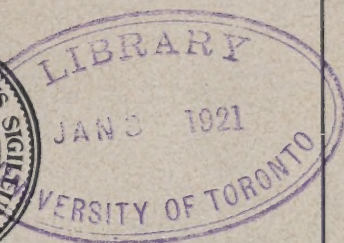


BULLETIN OF THE EXTENSION DIVISION, INDIANA UNIVERSITY

VOL. V

BLOOMINGTON, IND.

No. 1



Sixth Conference on Educational Measurements

SEPTEMBER, 1919

Entered as second-class mail matter October 15, 1915, at the postoffice at Bloomington, Indiana, under the Act of August 24, 1912. Published monthly by Indiana University, from the University Office, Bloomington, Indiana.

For Sale by the University Bookstore, Bloomington, Ind. Price 50 Cents.

A limited number of copies of this bulletin will be distributed free of charge to citizens of Indiana.

SIXTH ANNUAL CONFERENCE
ON
EDUCATIONAL MEASUREMENTS



Held at Indiana University, Bloomington, Ind.,
Friday and Saturday, April 18 and 19, 1919

PUBLISHED BY
THE EXTENSION DIVISION OF INDIANA UNIVERSITY
1919

Contents

	PAGE
INTRODUCTORY STATEMENT. By WALTER S. MONROE, Director, Bureau of Coöperative Research, Indiana University.....	3
DIAGNOSIS OF LANGUAGE ERRORS. By W. W. CHARTERS, Dean of the School of Education, University of Illinois.....	6
DIAGNOSIS OF GRAMMATICAL ERRORS. By W. W. CHARTERS.....	13
DEMONSTRATION OF THE USE OF GROUP TESTS FOR INTELLIGENCE. By SID- NEY L. PRESSEY, Research Assistant in Psychology, Indiana University.	25
A GROUP SCALE OF INTELLIGENCE FOR THE FIRST AND SECOND GRADES. By MRS. SIDNEY L. PRESSEY, Graduate Student, Department of Psychology, Indiana University.....	38
SCHOOL SURVEYS BY MEANS OF GROUP TESTS OF INTELLIGENCE. By SIDNEY L. PRESSEY.....	46
GROUP TESTS FOR INTELLIGENCE IN THE BEDFORD, INDIANA, SCHOOLS. By EMERY W. MONTGOMERY, Superintendent of Schools, Bedford, Indiana.	54
MENTAL DEFECTIVES IN THE RURAL SCHOOLS OF COUNTY H IN RELATION TO THEIR PHYSICAL ENVIRONMENT. By HAZEL HANSFORD, Research Fellow in Psychology and Sociology, Indiana University.....	56
SOME EXPERIMENTS WITH MENTAL TESTS AS AN AID IN THE SELECTION AND PLACEMENT OF CLERICAL WORKERS IN A LARGE FACTORY. By Chester S. CARNEY, Consultant with Scovell, Wellington, and Company, Indus- trial Engineers, Boston, Mass.....	60
INTELLIGENCE TEST OF HIGH SCHOOL SENIORS. By OSCAR H. WILLIAMS, State High School Inspector, Indianapolis.....	75
SCIENTIFIC CURRICULUM CONSTRUCTION. By W. W. CHARTERS.....	78
NEXT STEPS IN EDUCATIONAL MEASUREMENTS. By WALTER S. MONROE....	94
EDUCATIONAL SERVICE IN IOWA. By ERNEST J. ASHBAUGH, Director, Bureau of Educational Service, Extension Division, University of Iowa.....	104
SOME RECENT DEVELOPMENTS IN SPELLING. By ERNEST J. ASHBAUGH.....	113

Introductory Statement

THE Sixth Annual Conference on Educational Measurements was held under the auspices of the School of Education at Indiana University, Friday and Saturday, April 18 and 19, 1919. The attendance of superintendents and teachers was very good, and unusual interest was indicated in all sessions of the Conference. This Conference is held annually by the School of Education as one means for bringing to the attention of superintendents and teachers of the state significant matters relating to educational progress. It is gratifying to the School of Education that it is evident that the school men of the state have come to look upon this Conference as one of the most important educational gatherings of the year.

The program of the sessions is given below. This year no stenographic record was obtained of the Conference, and the published proceedings include only the manuscripts submitted by the speakers on the program. In a few cases no manuscript was submitted because conditions unforeseen at the time the program was published made impossible a satisfactory report.

FRIDAY, APRIL 18, 1919

Morning Session, 10:45 a.m.

Presiding, Linnaeus N. Hines, State Superintendent of Public Instruction, Indianapolis.

- 10:45. Opening Address: William Lowe Bryan, President of Indiana University.
- 11:00-11:45. Address: Diagnosis of Language Errors. W. W. Charters, Dean of the School of Education, University of Illinois.
- 11:45-11:55. Report: The Progress and Promotion of Pupils in Certain Indiana Cities. Walter S. Monroe, Director of the Bureau of Coöperative Research, Indiana University.

Afternoon Session, 1:15 p.m.

Presiding, Walter S. Monroe.

- 1:15. Demonstration of the Use of Group Tests for Intelligence. Sidney L. Pressey, Research Assistant in Psychology, Indiana University.
- 2:15-4:00. Mental Measurements in the Army and Public Schools.
1. Brief Reports on Results of Group Testing in the Army and in Indiana Public Schools. William F. Book, Professor of Educational Psychology, Indiana University; Oscar H. Williams, State High School In-

spector, Indianapolis; Rudolph A. Acher, Department of Psychology, State Normal School; Emery W. Montgomery, Superintendent of Public Schools, Bedford; A. H. Douglas, Superintendent of Schools, Logansport; Earl E. Ramsey, Superintendent of Schools, Bloomington.

2. Mental Defectives in the Rural Schools of County H. in Relation to their Physical Environment. Hazel Hansford, Research Fellow in Psychology and Sociology, Indiana University.
 3. A Group Scale of Intelligence for the First and Second Grades. Mrs. Sidney L. Pressey, Graduate Student, Department of Philosophy, Indiana University.
 4. School Surveys by means of Group Tests of Intelligence. Sidney L. Pressey.
- 4:00-4:45. Address: Educational Service in Iowa. Ernest J. Ashbaugh, Director of Educational Service of the Extension Division, University of Iowa.

Evening Session, 8:15 p.m.

Presiding, Acting Dean William W. Black, School of Education, Indiana University.

Address: Scientific Curriculum Construction, W. W. Charters.

Address: Next Steps in Educational Measurements. Walter S. Monroe.

SATURDAY, APRIL 19, 1919

Morning Session, 9 a.m.

Presiding, Earl E. Ramsey, Superintendent of Schools, Bloomington.

9:00-9:45. Report: Some Recent Developments in Spelling. Ernest J. Ashbaugh.

9:55-10:40. Report: Some Experiments with Mental Tests as an Aid in the Selection and Placement of Clerical Workers in a large Factory. Chester S. Carney, Consultant with Scovell, Wellington, and Company, Industrial Engineers, Boston, Mass.

Discussion.

10:50-11:30. Round Table Discussion: Educational Diagnosis and Corrective Instruction. W. W. Charters.

11:30-11:50. Presentation of the Plans of the Bureau of Coöperative Research. Walter S. Monroe.

Afternoon Session, 1:30 p.m.

Presiding, Hubert G. Childs, Professor of Education, Indiana University.

1:30-2:15. Diagnosis of Grammatical Errors. W. W. Charters.

2:20-3:00. Report: Pupils' Errors in Arithmetic. Walter S. Monroe.
Discussion.

The keynote of the meeting was a plea for diagnosis and more complete interpretation of educational measurements in order that the use of educational tests may result in the greatest possible improvement of instruction. This topic was emphasized by Dean Charters of the University of Illinois in two addresses and was mentioned incidentally a number of other times. It is clearly a next step. It is practically impossible to justify the use of educational tests unless it results in improving instruction. Up to the present time we have been primarily concerned with getting acquainted with educational tests and learning how to use them to measure the ability of pupils. We are now ready to take up the use of the results of the tests in improving instruction.

A very pleasing and satisfying event which is not listed upon the program was the Teachers' Dinner held at the University Commons on Friday evening at 6:30. Acting Dean Black of the School of Education presided at a brief program of after-dinner speaking in which technical matters relating to education were almost entirely eliminated.

WALTER S. MONROE,

Director, Bureau of Coöperative Research, Indiana University.

Diagnosis of Language Errors

W. W. CHARTERS, Dean of the School of Education, University of Illinois, Urbana, Ill.

THERE is a common feeling that attention to errors and to defects in language produces average ability to write and speak, but does not promote artistic speech. It is commonly felt that this attention to errors brings the individual student up to the average level but does not incite him to superior ability.

This opinion is incorrect. It presupposes that the standard in diagnosis of errors is the average, and it takes no more than a mere mention of the fact to make it clear that the standards can just as well be those of the writings of Stevenson, Shakespeare, or Lincoln, and that deficiencies may be gauged by the best as well as any other standards including the average. Diagnosis is not concerned merely with errors such as *I seen*, and *It ain't*, but it is also concerned with the superiority of *Half a league, half a league, half a league onward*, over the mathematically accurate and shorter *A league and a half onward*. While it seeks to correct *He don't*, it notes with equal clearness the inferiority of *She is a peach*, to *She walks in beauty like the night of cloudless climes and starry skies*.

There is, to be sure, a literary spontaneity which of its own accord rises to different heights of excellence, partly by virtue of native talent and partly by undirected study. Sometimes this rises to spectacular levels as in the case of Burns but usually it is on the plane of the average. The great mass of school children, however, will, at best, do well if they speak correctly, and only a small number may hope to attain to brilliance of expression.

But the brilliant and the average alike can be improved by attention to their deficiencies. If standards are set on the lower level of correct speech, artistry will suffer, but if the standards established by the masters is set, artistry will be developed in those capable of attaining unusual facility of expression. Diagnosis of difficulties is the most satisfactory method of improvement for both the average and the brilliant student.

Presumably a teacher can read the oral and written language of children appreciatively and may thereby look for excellent expression and ignore mediocrity. He may stress the fine and striking words and phrases and this has its advantages in helping students to be original and brilliant. But he can also look for deficiencies and errors, note where language falls below a standard, and direct the writer to correct and improve. Instead of taking as his motto the idea, "This is fine", he may substitute, "This is wrong and can be improved." It is obvious that both of these can be used, but improvement is more dependent upon the notice of deficiencies than upon the notice of accomplished excellencies. An experienced writer asks the criticism of an experienced

critic in order to obtain some evaluation of good and bad, but particularly in order to be made aware of the bad. However, children more than adults need to have both excellence and defects pointed out because they lack confidence in their own powers and they become discouraged if only the defects are noticed.

The foregoing is stated merely for the purpose of reinforcing what was said above to the effect that while attention to excellence is valuable, improvement depends primarily upon attention to deficiencies.

It is obvious that standards are necessary bases for diagnosis. The rejects of the army could not have been determined without reference to military standards. The presence of flat feet was discovered only after the proportions of a perfect foot had been obtained, and undernourished bodies were determined only after the standards for a well-nourished body were known. An error in language is noticeable only when correct forms have been ascertained, and deficiencies in artistic language which is grammatically correct are apparent only to those who have standards of rhetorical excellence.

In the determination of standards in language some slight progress has been made in the scales of Hillegas, Ballou, and Willing, and in the Nassau County Supplement. These assume a standard of both rhetorical and grammatical excellence which is not confined to average performance. The method of determining which composition is good and which better is that of comparing one composition with another. These compositions were written by children, and theoretically the more carefully prepared of these scales begins with zero ability in composition, and on this basis by the method of equally observed differences, compositions are selected which are one unit, two units, etc., better than zero. It is evident, of course, that this method could have been continued by the inclusion of other compositions until Stevenson and Lincoln were rated in comparison with children's compositions and with zero ability in writing. The fact that the Hillegas scale has only ten steps between zero and the best is due entirely to the fact that the selection of his compositions limited him to that range. The inclusion of compositions superior to those he had would have made it possible to continue the process of adding steps like those of the ladder of the patriarch, Jacob, until the literary heaven of the great masters is reached.

But these composition scales just mentioned are not useful for diagnostic purposes. The reason for this is that, in general, they are based upon standards which are unanalyzed. For instance, Hillegas had his graders rate the compositions merely as better and poorer than each other and without regard to any standards other than those the graders had in their own minds. And these were very general standards even tho some graders laid great emphasis upon punctuation, others undue emphasis upon grammar, and still others valued artistic expression very highly. Even tho Ballou improved on this analysis by dividing the compositions into the four major divisions of narration, description, exposition, and argumentation, and Willing into the content and form, the same criticism holds. The scales are the result of the application of unanalyzed standards and are not, therefore, particularly useful for diagnostic purposes.

It should be interpolated at this point that a diagnostic test must be able to reveal specific difficulties since improvement is obtained by first establishing a standard and then locating the specific points at which the standard is not attained. For instance, the army has done its superficial duty when it merely rejects an applicant, but if this reject is turned over to an improvement agency it is necessary to do more than reaffirm the rejection of the army officials or to advise the reject to improve his health. A diagnosis for purposes of improvement must be much more detailed and must be carried on until the reject is found to have defective teeth, flat feet, a curved spine, or weak eyes. Only when physical diagnosis has been carried to this point is it useful for improvement, and in improvement of speech ability the same process holds. It is not sufficient to rate compositions, and, after affirming a grade to be 184, to return it to the pupil with the admonition to bring it up to 300 on the next attempt. Rather, the pupil must find in his returns specific points at which improvement is necessary. Once these have been determined, corrective methods may be introduced, but not until then, because improvement in reaching a standard is specific and not general. General improvement is in reality the sum of detailed improvements and is not independent of them.

To improve the Hillegas scale for diagnostic purposes it is necessary to analyze composition writing into its elements and obtain a scale for each.

It is a confession of the pioneer that in attempts to obtain diagnostic tests as opposed to general ability tests the only tests yet made have to do with correctness rather than with artistic excellence. A beginning had to be made somewhere, and it is easier to set up standards of accuracy than of artistic and rhetorical excellence.

Starch has made one such attempt in his punctuation scale. A series of unpunctuated sentences has been prepared and the pupils are requested to punctuate them correctly. This enables the teacher to determine not only which punctuation marks each pupil fails to use but also which errors are common to the greatest number of pupils and, therefore, need the greatest amount of immediate attention. It also determines the relative efficiency of classes and of individuals in each class.

The other attempt of a similar nature is that of the writer in a pronoun scale recently issued in preliminary form. At present several thousand copies of this scale have been printed and used and standards of performance are now being established and will soon be published.

For our purposes a brief description of the methods of determining the elements will be sufficient.

The purpose of the scale is frankly that of determining how far the pupils deviate from accuracy of grammatical form. No attempt is made to determine rhetorical or artistic ability.

The scale is a by-product of studies made in tabulating the grammatical errors of school children, principally in Kansas City and Detroit, but with verifying studies in other states, studies of sufficient extensity to show that the types of grammatical errors are common to all parts of the nation and that the differences in speech between one section and

another are largely differences of vocabulary, pronunciation, enunciation, and non-grammatical idioms rather than differences in grammar.

In all, some twenty-five thousand oral errors were tabulated and classified. Fortunately, it was found that these twenty-five thousand cases could be classified under twenty-one heads. For instance, it was found in Detroit that fourteen subheads under the twenty-one rules comprised 6,500 of the 11,200 errors, or nearly 60 per cent of the cases. These were, in order of frequency: *ain't got, give for gave, was for were, seen for saw, can for may, ain't for isn't and aren't, done for did, John he, is for are, don't no* (double negative), *them for these and those, me and Catherine did, come for came, and don't for doesn't.*

In this collection of errors all types were included and it was a simple matter to make selections of errors and to present them to pupils. The collection provided further for the making of equivalent forms of the same test so that children can be tested out a second time with similar material upon identical points.

I have distributed the language scale¹ among you and an examination will show two items of importance. The first is the value attributed to each sentence. These were determined by the usual method of finding among 3,500 children how many failed to write them correctly and of rating these in terms of sigma of the normal probability curve. This means that on the basis of difficulty an item with "3" suffixed is three times as difficult as one with "1", and one and one-half times as difficult as one with "2" after it. This is not exactly accurate, as, obviously, the decimal has been dropped and the nearest integer substituted. Furthermore, in our next form of the scale we shall drop the weights and give them all equal value because we have found that the correlation between scores determined, on the one hand, by adding weighted suffixes, and, on the other hand, by giving a value of one to each sentence, is 97, which is so large as to make the extra labor of weighting unnecessary.

The second point is the method of examination. It is to be noted that the preliminary statement says, "Some of these sentences are right and some are wrong." This is significant because the child approaches each sentence in the natural way. By this I mean that he has no more hint as to which sentence is right and which is wrong than he has when he prepares to use the sentence in practical speech. This is an improvement over the plan of having him choose alternative forms, as in (Who, Whom) do you see? where he has the two forms before him in parentheses and is asked to cross out the wrong form. Our plan is better because the presence of *who* and *whom* makes him more alert mentally than he would be if the sentence were given him with an innocent-appearing error quietly lodged within it. The labor of writing is greater and the task of marking is somewhat longer, but the returns constitute a much more accurate description of what children actually do.

A portion of this language scale is reproduced herewith. The remaining items of the scale may be obtained by referring to the Diagnostic Language and Grammar Test reproduced on pp. 16-22. The distinction between this test and the one reproduced here is that it has spaces added for giving the grammatical reason for the correction. The type of this test as well as the other test reproduced in this volume has been reduced somewhat in size in order to facilitate the reproduction.

UNIVERSITY OF ILLINOIS

Urbana, Illinois

BUREAU OF EDUCATIONAL RESEARCH

DIAGNOSTIC LANGUAGE TEST

PRONOUNS

By W. W. CHAMBERS

Grades III to VIII

County

State

Date

Pupil's Name

Age

Grade

DIRECTIONS FOR GIVING THE TEST

After telling the children not to read the papers, ask those on the front seats to distribute the papers, placing one upon the desk of each pupil in the class. Have each pupil fill in the blanks at the top. Then read the following instructions with the pupils.

Instructions to be read by teacher and pupils together

"This test is given to pupils who have studied language lessons to see how well they are able to tell when sentences are right and when they are wrong. In these pages will be found exercises like the following:

1. I told him to go

"The plan is to read this sentence over carefully and see if it is correct. If it is correct copy the sentence exactly on the dotted line below the sentence. The sentence, *I told him to go*, is correct so we shall write it on the dotted line below. We shall now write this sentence in." (The teacher should pause to allow time for doing this).

"If the sentence is not correct we are to write the correct form on the dotted line below. Let us try one that is not correct."

Pupil's Score

2. I told she to go.

"The correct form is *I told her to go*. So we shall write *I told her to go* on the line below." (The teacher should pause to allow time for doing this.)

"When you have copied all the sentences on the first page you will stop work."

The teacher should pass around the room and see that all the children fill the above forms correctly and when all have finished he should say, "Begin work on the first page below." This is not a speed test so ample time should be allowed for each to finish. The test should be completed in two periods taking the first page in the first period. Collect a pupil's paper as soon as he has finished. At the second period have the papers distributed giving to each pupil his paper. As at the first period tell them not to read the papers until told to do so and re-read the instructions with them.

	Value 1	7. Who do you want?	Value 2
1. John went to town.			
2. May Inez and me go?	1	8. Him and I will do it.	2
3. It teaches a person something you may use.	3	9. Us boys did it.	2
4. Your girls plays too hard.	1	10. The boy was raised by it's own mother.	2
5. When one lives in town they hear noises.	3	11. These books are young.	1
6. That is hern.	2	12. It was a book with riddles in them.	1

I may say, parenthetically, that to avoid the labor of writing we are now experimenting with directions to see if we cannot have the pupil merely write, on the line below, the correct form of the incorrect word in the sentence above, and neglect the other words.

It is obvious that no time limit can be set for such an exercise since speed in writing would be a determining factor of large influence. On the other hand, in so far as the test is diagnostic, the teacher is more concerned in finding what the pupil can do with these errors when he is given the opportunity to attack each than he is in seeing how many he can do in a given period.

A final word needs to be said as to the use of these tests in diagnosis.

This language test gives examples of all types of errors which were found in pronouns. When a test has been given, the teacher has before him a fairly complete analysis of the pupils' ability to recognize the correct forms. It should be pointed out, of course, that such would not be the case in errors in verbs because while we can mathematically determine all possible errors in pronouns, since there are such a small number of them, it would be an almost impossible task to determine all verbs in which errors might presumably be made. But in either case, if frequency is made the basis of selection, a fairly complete list of those errors which the class make can be determined.

Once these errors have been ascertained it is then possible to find out the errors which are most frequent in the class as a whole and thereby it is quite easy to arrange to lay particular stress upon the correction of the most common ones. Of equal importance is the fact that the teacher and each pupil may become aware of the special difficulties of each pupil and thereby incite him to work harder in correcting his difficulties.

If, however, the study stops at this point, little has been accomplished. In other words, as time goes on it is necessary to tie up with the diagnosed difficulties the corrective measures and these are of at least three sorts. In the first place, language games can be given to fix the correct forms in the minds of the children. Many of these games are available in certain books prepared for that purpose and other games are in current use among teachers. In the second place, language exercises such as those found in blanks, alternative forms, etc., are excellent. And, in the third place, eternal vigilance in correcting every mistake whenever it occurs is an absolute essential. And underlying all of these must lie the desire within the pupil to correct his speech. The absence of this desire is the most difficult situation that has to be faced and it would be well some day for a diagnostic test to be made to determine the amount of desire and the points at which the desire is not present.

In general, then, the situation with regard to diagnosis of linguistic errors is this: Some standard tests of a general sort have been made which have little diagnostic value. The only attempts to obtain diagnostic tests have to do with correctness and have nothing to do with rhetorical excellence other than that which resides in, or comes from, accuracy of statement, but it is comparatively easy to prepare such diagnostic tests by an analysis of the elements of artistic excellence and this, it is to be hoped, will be the next line of study which will be taken up by those interested in the subject.

Diagnosis of Grammatical Errors

W. W. CHARTERS

THE primary function of grammar is that of analyzing the sentence into its elements and showing the relation of each element to the other. This analysis reveals the seven parts of speech and their declension forms, phrases and clauses, and the syntactical relations such as agreement and modification.

While the function of grammar is thus freely acknowledged by all, the function of its study in the grades has not been so universally agreed upon. Historically, its study has been justified for many years upon the basis either of knowledge for its own sake, or of mental training. In the former case, it is claimed that the sentence is such an important social instrument, since one uses sentences more frequently than any other tool made by man, that a knowledge of its structure is an essential part of the education of anyone mature enough to understand it. In the latter case, it has been claimed that the study of grammar trains the powers of reasoning and discrimination in an unusually efficient manner and the assumption has been made that this training can be carried over to other situations in which reasoning and discrimination are useful. In recent years certain investigations and discussions have seemed to discredit this assumption, tho to what extent has not yet been fully determined.

But whatever the real extent of the transfer of the training, the startling claims of the opponents of formal discipline have caused such a thoroughgoing re-examination of the function of the study of grammar in the grades that as a result a third function is now being widely discussed and elaborately analyzed. This is the idea that the justification of grammar in the grades lies in its use for the correcting of errors. The claim is made and substantiated that a knowledge of the reasons for correct forms aids the doubtful memory of forms. It can be pointed out, for instance, that a child that has been using *him* and *me* in the nominative and has been taught to use *he* and *I* becomes confused after a time and is uncertain as to which form is correct. Accordingly he sometimes resorts to the use of *he* and *I* exclusively in both the nominative and objective case. At other times, by a sort of random selection, he uses them indiscriminately. It is at this point that a knowledge of nominative and objective cases is useful when their object and subject relations give him an anchorage by which to stabilize the use of both forms.

Whether or not this function of grammatical study is the correct one we are as yet unable to say. But it seems so reasonable and is so rich in suggestion that investigators are amply justified in working the idea thru and developing all the value contained therein. I, personally, am quite firmly of the opinion that unless this function can be justified,

there is no other adequate justification of any sort for the spending of two years or more in the grades upon the subject. My reasons for thinking this are the following:

In the first place, the subject is barely within the range of ability of seventh- and eighth-grade children. Studies of children's ability to parse show a much lower level than is found in spelling or arithmetic, reading, geography, or history. Grammar is just on the margin of inclusion, as shown by the further fact that an excellent teacher of grammar can teach a majority of her class, but by no means all, to parse and analyze; while an average teacher makes a failure of the task. Because of these facts it would appear that the first reason for the study just mentioned, namely, that children should know about such an important social tool, is not valid for seventh- and eighth-grade children and that the subject should on those grounds be left for study in the high school. Or, if this were not done, then a less thoroughgoing study should be made of the structure of the sentence if it is taught in the seventh and eighth grades. Such a study would have as its function the teaching of the parts of speech, declension forms, and some analysis of simple sentences together with the recognition of phrases and clauses. This study should resemble our present study of grammar to about the same degree that hygiene and physiology resemble scientific college physiology or nature study resembles high-school physics and biology.

In the second place, the evidence against the transfer of training in reasoning and discrimination has been so severely shattered that it is not likely, even under the most favorable construction, to maintain sufficient validity to warrant the teaching of grammar, especially since it is so difficult for children. Because of this fact, I do not see how we can justify the inclusion of the subject in the seventh and eighth grades upon the grounds of its training value. Even if there were considerable transfer of training, so few children master analysis and parsing adequately that in most grammar classes so little is gained in the way of training that there is nothing of importance to be transferred.

In the third place, it remains to be seen whether transfer will function in the correction of errors. Theoretically, it seems that such would be the case since reasons for facts are supposed to aid the memory of facts. Logical memory is usually more retentive than mechanical memory, but whether the subject is too difficult to be used effectively even for this very valuable purpose, and whether the effort put forth pays in results, remains to be determined. However, the use of grammar in correcting errors does not require a full analysis and parsing of all sentences. In many cases certain scraps of information are usually sufficient, as, for instance, in the sentence (Who or whom) do you see? Here it is necessary to parse only one word, *whom*, and even in that case all that is necessary is to know objective case and object of the verb *do see*. The other portions of a complete parsing of the pronoun are not necessary nor does the parsing of the remainder of the sentence need to be carried on. Then, too, pronouns and past tenses constitute most of the difficult grammatical facts since there are few errors in nouns, prepositions, conjunctions, and interjections, and the errors in adjectives and adverbs are comparatively simple to understand. But whether these

few facts can be used without a knowledge of more grammar is as yet unknown.

Consequently, because of the difficulty of grammar for the seventh and eighth grades and the lack of knowledge concerning transfer, the inclusion of formal grammar for elementary schools rests upon its relation to the correcting of errors, or upon nothing.

If we accept this functional point of view of the study of grammar, two or three considerations follow:

In the first place, the grammar to be taught must be based upon the grammatical errors which children make. These are of two sorts, oral and written. Both types include errors not usually classed as grammatical, such as the spelling of *to*, *two*, and *too*, which can be explained on the basis of *to*, a preposition, *two* as an adjective, and *too* as an adverb, or the punctuation of a sentence, as the use of commas before and after nouns in apposition. Written grammatical errors include certain items which do not appear in oral speech, such as the spelling of homonyms, as above, capitalization of proper nouns, periods, exclamation marks, etc. But these errors, both oral and written, can be determined, and in the case of oral errors, particularly, quite extensive studies have been made.

Once these errors have been found they can be classified, and certain rules, about thirty in all, are found to be broken. Then it becomes necessary to find out what additional facts of grammar must be known in order that the rules may be understood. For instance, the rule which runs, The subject of a verb is in the nominative case, necessitates for a clear understanding a knowledge of nouns, pronouns, verbs, subject, and predicate. When the rules that have been broken and the grammatical declensions which are necessary to understand them have been collected, we have the course of study in grammar much in the form worked out in the Kansas City Grammar Study and reported in Part I of the Sixteenth *Yearbook* of the National Society for the Study of Education.

Quite significant is the fact there noted that mood, infinitives, gerunds, and participial constructions are excluded from the course of study and, on the whole, a much simpler course of study is obtained than is found in the typical modern textbook; but whether the analysis of sentence in terms of this simpler subject-matter is too difficult for seventh- and eighth-grade children remains to be seen.

In addition to the fact that the subject-matter must be determined upon the basis of errors and to the further fact that grammar is used for the correcting of errors, the hold upon the subject must be automatic. Automatic oral parsing is an absolute essential to efficiency. Greater speed in parsing is necessary than in the fundamental processes of arithmetic, for while arithmetic is used only occasionally, the necessity for choosing between right and wrong forms in grammar is constant. Whenever a person talks he needs the information. Furthermore, he has to make his decisions on the spot. If, as he converses, he forgets whether *He doesn't* or *He don't* is correct and calls upon his knowledge of grammar to assist, he must make his decision about singular noun and singular verb forms instantaneously. If he cannot do it

UNIVERSITY OF ILLINOIS
Urbana, Illinois

BUREAU OF EDUCATIONAL RESEARCH
DIAGNOSTIC LANGUAGE AND GRAMMAR TEST

PRONOUNS

By W. W. CHARTERS

Grades VII and VIII

PART I

Post Office..... County..... State..... Date.....
School..... Pupil's Name..... Age..... Grade.....

DIRECTIONS FOR GIVING THE TEST

After telling the children not to turn the papers over, distribute Part I, this side up. Have each pupil fill in the blanks at the top. Then read the following instructions with the pupils:

Instructions to be read by teacher and pupils together

"This test is given to pupils who have studied language and grammar to see how much they know about pronouns. On the next page will be found exercises like the following:

1. I told him to go.

"Read this sentence over carefully and see if it is correct. If it is correct, copy it exactly on the dotted line below. The sentence, *I told him to go*, is correct so we shall write it on the dotted line below. We shall now write this sentence in." (The teacher should pause while this is done.)

Pupil's Scores

Language.....

Grammar.....

"If the sentence is not correct, we are to write the correct form on the dotted line below and on the right half of the page we shall give the reason for making the correction. Let us try one."

2. I told he to go.

"The correct form is, *I told him to go*, so we shall write *I told him to go* on the dotted line below." (The teacher should pause while this is done.)

"On the right-hand side we are to tell why we make the change on the left-hand side. We always begin our reason for the change with the word *Because*. In this case we write—*Because him is in the objective case, object of the verb, told*. (The teacher should pause until the pupils write the reason at the right of the sentence.) We must not write too large and we may use both of the lines on the right-hand side so that we may have plenty of room for giving our reasons."

"We shall now try a third exercise."

3. It is him.

"The correct form is *It is his*. We make the change *Because his is the possessive form of the pronoun, he*. Fill this in correctly, writing the correct form on the line below and beginning the reason on the right-hand side with the word *Because*. Do this now.

"Remember that if the sentence on the left is correct, we do not need to write anything on the right-hand side because we have to make no change and so have to give no reason for making a change."

"The teacher should pass around the room and see that the children fill these in correctly and when all have finished he should say, "Turn your papers over and begin the work." This is not a speed test so ample time should be allowed to each to finish. This test should be completed in two periods during the same day. Part I should occupy the first period, and Part II the second. At the beginning of the second period, go over the above instructions so far as may be necessary and distribute Part II. After children have finished Part II, collect papers and assemble Parts I and II for each child, fastening them together.

Grammatical Value	1	2	3	3	3	3	3	3
1. John went to town.	1							
2. May Inez and me go?	1	2						
3. It teaches a person something you may use.	3	3						
4. Your girls play too hard.	1	4						
5. When one lives in town they hear noises.	3	5						
6. That is here.	2	6						
7. Who do you want?	2	7						
8. Him and I will do it.	2	8						
9. Us boys did it.	2	9						

10. The boy was raised by its own mother.	2	10	3
11. These books are yours.	1	11	3
12. It was a book with riddles in them.	1	12	3
13. Are those them?	3	13	3
14. Mary and her mother are here.	1	14	1
15. It was only us.	3	15	3
16. They made baskets and filled it with holly.	1	16	3
17. They went in a ship who sailed on Friday.	2	17	3
18. The pencil is his.	1	18	2

PART II

Pupil's Name	Language Value	Date	Grammar Value
19. I am older than him.	2	19	3
...			
20. I and my sister went home.	2	20	4
...			
21. Them are mine.	2	21	3
...			
22. He pushed John and I.	3	22	3
...			
23. He corrected his work his self.	1	23	3
...			
24. It was me.	2	24	3
...			
25. He struck a girl which was near by.	3	25	3
...			
26. Annie called to you and I.	2	26	3
...			

27. Her and I used to play together.	2	27	2
28. Was that him?	3	28	3
29. I want him to go.	1	29	1
30. That's her.	3	30	3
31. You and them must decide.	2	31	3
32. Every girl took their seat.	2	32	3
33. We found ourself in a large room.	2	33	3
34. They kept themselves hidden.	2	34	3
35. Who did you speak to?	2	35	3

36. He found a hole what he could get into.	2	36	4
37. They were afraid to show themselves.	2	37	33
38. The princess called he Moses.	1	38	34
39. The boys will soon catch they.	1	39	34
40. It is whom?	3	40	33
41. Brother John and myself went home.	2	41	33
42. I told she to be still.	1	42	34

at once he is compelled to use whichever comes first, if he happens to be talking enthusiastically; or, he is compelled to halt while he is deciding the matter and thereby loses greatly in effectiveness of speech. In arithmetic, if it becomes necessary, he can take pencil and paper and figure the result, but in speech he cannot stop and diagram his sentences.

From this point of view, written analysis and written parsing are only means of attaining facility of oral parsing. They cannot be the end of effort.

Perhaps when this is recognized and grammar is seen in relation to its automatic outcome and all drill devices are used by teachers to attain this end, grammar will be more useful than it now seems to be as an aid in speech. Teachers have taught grammar for so many reasons other than the correction of errors that they have not made drill sufficiently rapid to function properly and this may be the full cause of the failure of functioning. For if adequate speed is not acquired, less than adequate speed is useless. It is a case of all or nothing.

With these two standards once recognized, namely, a grammar based on errors and a study of grammar culminating in automatic action, diagnosis may begin. Complete diagnosis will be of two sorts. These will be the diagnosis of the ability of children to apply rules when time is given, and of their ability to do it speedily. For only the first of these have tests been made, if we do not count Starch's tests to determine pupils' ability to recognize parts of speech and declension forms. In the grammar test on pronouns¹, it will be noted that it consists in part of the language test which was distributed yesterday. The items of the language test are shown on the left-hand side of the page. In addition, on the right are blanks for giving the reasons for the corrections made on the left.

The weighting should also be noted, for the same observations as to method of determining these and the expectation of discarding them in our next form, as were made yesterday, hold.

When we consider the decision as to whether or not the grammatical reason given in any case is adequate, the task is much more difficult than in the case of language. The query as to what is the correct answer is difficult to decide upon, so considerable help in making the decision is necessary.

Negatively, it is obvious that such answers as, *It is not polite*, or *It does not sound right*, are not correct since grammatical reasons must be stated in terms of grammatical facts and rules. These are merely language reasons.

Positively, an answer is adequate if the declension and the syntax are both given. In the sentence, *Was that him?*, *Because "he" is a pronoun in the nominative case, complement of the verb "was"*, is adequate since it mentions both the case (nominative) and the syntactical relation (complement of the verb *was*). If all answers included these two factors, the decision would be easy, but what shall we do with an answer such as this: *Because "he" is nominative case?* Is this adequate? It is not for a teacher of grammar who will insist upon the statement

¹Reproduced on the preceding pages.

of relation, but for the practical purpose of seeing if the pupil has some grammatical reason for his correction it is probably sufficient. To know that the nominative case should be used is adequate, particularly when the student may know the syntactical relation but has not been drilled in the precise form which the answer should take in such a test as this. Again, a pupil may say, *Because "he" is the complement of the verb "was"*. Such an individual, of course, omits mention of case, but he may have forgotten to mention the case which may have been well known to him and, therefore, for practical purposes of speech the statement of relation may be a sufficient answer. As said above, the teacher will not be satisfied with the inclusion of one of these elements, but the grader, it seems to me, should give the pupil the benefit of the doubt, until such time, at least, as definite forms of answering the test have been agreed upon and are commonly used.

It would seem, therefore, that for testing purposes either declension forms and syntactical relations, or declension forms *or* syntactical relations alone would be considered an adequate answer. But for the diagnostic purposes of the teacher both should be included.

What was said with regard to language diagnosis holds for the diagnosis of grammar. Diagnostic tests must show specific difficulties and not general deficiencies. The teacher should be able to know after the test is given which rules the children can use when given sufficient time to write them and which children are better and which are poorer, and in which respect each child is superior or deficient.

The diagnostic test in grammar obviously cannot be given until after the topics in grammar have been studied because the children do not know the terminology before that time. But once they have been studied it is then possible to give the test to succeeding classes thruout the grades and the high school from year to year to determine what points still need to be drilled upon by succeeding English teachers.

As a matter of practical importance it must be stated that any superintendent can construct diagnostic tests in language and grammar for himself. If he wishes to compare his classes with those of other systems he should use the standard forms, but for purposes of diagnosis within his own system he can make his own tests and even have them written on the blackboard. This is the case because such tests do not include a time element and, furthermore, our studies show that weighting is not necessary. If each sentence is given the value of *one*, the results are almost identical with those in which weighting values are used.

Consequently, what such a superintendent needs to do is, first of all, to collect the errors made in his system and from these select certain important ones in the form of incorrect sentences. Then he should ask the teachers to have the children write the corrections, in the case of language diagnosis, and to give the reasons for corrections, in the case of grammar diagnosis.

In the matter of grammar correctives it may be stated that these may be collected, but chief reliance must be placed upon definitions of grammar and parsing and analysis, particularly as an oral exercise, until the children can instantaneously tell, with reasons, which are the correct and which the incorrect forms.

Demonstration of the Use of Group Tests of Intelligence

Reported by SIDNEY L. PRESSEY, Research Assistant in Psychology,
Indiana University

At 1:15 on Friday afternoon Dr. Sidney L. Pressey gave a demonstration of the "Cross-out" Tests.¹ He tested the audience, using a modified form of the directions, as given below:

TEST I

Each of the twenty-five lists on the page is a sentence with the words mixed up. But in each list there is one, and only one, word that cannot be used in making the sentence. You are to cross out this extra word. Study the examples and then begin. (3 minutes.)

TEST II

In each of the twenty-five lists there are four things that are, in some important way, alike, and one thing that is different from *all* the other four. You are to cross this one thing out. Study the examples, and then begin. (2 minutes.)

TEST III

In each of the twenty-five lists the numbers are arranged according to some rule, but there is in each list one, and only one, number which breaks the rule. You are to cross out this number. Study the examples and then begin. (3 minutes.)

TEST IV

In each of the twenty-five lists you are to cross out the thing that is worst. Study the examples and then begin. (2 minutes.)

Immediately following this demonstration a class of 1A children were brought into the room and examined by Mrs. Pressey with the Primer Scale.² The directions and examination blanks are reproduced below:

¹This test is reproduced in full on pp. 26 et seq. In the blanks used at the Conference the correct words or numbers in all the examples were crossed out.

²This scale is reproduced in slightly reduced form on pp. 30 et seq.

Blank 1

MENTAL SURVEY SCALES

Schedule E

"Cross-out" Tests

Name Sex

Age Birthday Grade

Teacher School

Place State Date

TEST I. VERBAL INGENUITY

Examples: (a) See a I man on.
 (b) knife chair the sharp is.
 (c) in winter with covered trees is ground the snow.

- | | |
|--|-----------------|
| (1) the cat at see. | Score (1) |
| (2) boy was sky the sick. | |
| (3) bread sweep will the kitchen I. | (4) |
| (4) are going yesterday to-morrow we. | |
| (5) me mine give my straw hat. | |
| (6) brown the horse come is. | |
| (7) my suit dollars wear twenty cost new. | |
| (8) know ice big boys how skate to. | |
| (9) their soldiers for fight gun country. | |
| (10) teacher me from gave a pencil my. | |
| (11) brother lamp is my than I older much. | |
| (12) dusty road the is hot and miles. | |
| (13) in the chalk he brightest is boy class our. | |
| (14) roads hard to is climb steep very the hill. | |
| (15) broke his robin the flew little poor wing. | |
| (16) gave me candy brother my of knife a box. | |
| (17) told girl I I the to would her with home walk. | |
| (18) watch summer the man stole is jail who the in. | |
| (19) man whom the hat saw is you uncle my me with. | |
| (20) I not Monday do to bag like go to school on. | |
| (21) do not boy the I like who me school in sits desk behind. | |
| (22) old back only the chair legs has three black. | |
| (23) the from flood came mountain the valley bridge the down. | |
| (24) who the sing birds so flown during the sweetly to have south. | |
| (25) boy gold who highest watch jumps over the get the will a. | |

INDIANA UNIVERSITY
 DEPARTMENT OF PSYCHOLOGY

TEST II. LOGICAL JUDGMENT

Examples: (a) dog, cow, horse, oak, cat.

(b) book, newspaper, magazine, pamphlet, photograph.

(c) base-ball, skating, checkers, tennis, dominos.

- (1) coat, shoes, hat, gloves, sail.
- (2) oats, wheat, barley, cotton, rye.
- (3) satin, silk, fur, calico, serge.
- (4) fly, ant, bee, grasshopper, mouse.
- (5) bread, meat, vegetables, hay, fish.
- (6) river, lake, brook, pond, ocean.
- (7) beef, lamb, bread, mutton, pork.
- (8) house, barn, shed, cave, church.
- (9) sparrow, eagle, robin, owl, hawk.
- (10) tyranny, equality, democracy, freedom, liberty.
- (11) red, green, yellow, white, blue.
- (12) orange, tomato, apple, peach, plum.
- (13) mother, cousin, uncle, father, aunt.
- (14) oblong, triangle, square, octagon, circle.
- (15) iron, gold, platinum, copper, sulphur.
- (16) auto, horse, bicycle, steamboat, canoe.
- (17) grocer, lawyer, doctor, minister, professor.
- (18) mumps, deafness, tuberculosis, measles, diphtheria.
- (19) guitar, banjo, violin, flute, mandolin.
- (20) fox, wolf, dog, eagle, hawk.
- (21) kerosene, alcohol, coke, cinders, coal.
- (22) piston, boiler, valve, fire-box, dynamo.
- (23) pine, hemlock, maple, fir, spruce.
- (24) mayor, lawyer, alderman, senator, governor.
- (25) carpenter, bricklayer, machinist, bookkeeper, plumber.

TEST III. ARITHMETICAL INGENUITY

Examples: (a) 2 4 6 8 9 10 12

(b) 6 7 5 4 3 2 1

(c) 1 2 4 8 16 17

(1) 1 2 3 9 4 5

(2) 2 4 6 7 8

(3) 10 9 8 7 6 4

(4) 11 10 8 6 4 2

(5) 5 7 10 15 20 25

(6) 3 6 9 11 12 15

(7) 19 18 17 16 13 15

(8) 4 8 10 12 16 20

(9) 2 4 8 10 16 32

(10) 1 5 9 11 13

(11) 27 24 21 19 18 15 12

Score (2).....

(12) 1 3 5 7 9 10 11

(3).....

(13) 16 8 4 3 2

(14) 4 9 14 19 24 29 33

(15) 2 6 12 24

(16) 17 13 9 7 5 1

(17) 36 18 9 3

(18) 3 9 12 15 21 27

(19) 2 4 8 16 24

(20) 27 22 17 14 12 7

(21) 3 9 27 54 81

(22) 3 6 9 12 24 48

(23) 84 77 70 65 63 56

(24) 1 3 9 18 27

(25) 72 36 18 9 6

TEST IV. MORAL JUDGMENT

Examples: (a) Gambling, lying, drunkenness, murder, smoking.

(b) begging, gambling, trading, cheating, borrowing.

(c) dullness, foolishness, laziness, weakness, poverty.

- (1) fighting, killing, hating, quarreling, hurting.
- (2) borrowing, gambling, overcharging, stealing, begging.
- (3) love, hate, fondness, dislike, liking.
- (4) dancing, drunkenness, flirting, over-eating, smoking.
- (5) holiness, reverence, piety, obedience, wickedness.
- (6) kindness, roughness, cruelty, pity, harshness.
- (7) sincerity, honesty, self-confidence, frankness, deceit.
- (8) stinginess, carefulness, generosity, charity, economy.
- (9) patience, caution, eagerness, recklessness, thoughtfulness.
- (10) bashfulness, meekness, cowardice, foolishness, timidity.
- (11) justice, fairness, shrewdness, honesty, trickery.
- (12) laziness, idleness, tardiness, slowness, haste.
- (13) flattery, lying, fibbing, slang, gossip.
- (14) stupidity, dullness, foolishness, dishonesty, ignorance.
- (15) meekness, vanity, self-confidence, self-esteem, self-respect.
- (16) rudeness, impudence, insult, discourtesy, impoliteness.
- (17) patriotism, disloyalty, double-dealing, treason, fraud.
- (18) bigamy, flirting, indecency, immodesty, indelicacy.
- (19) gentleness, kindness, hesitancy, weakness, caution.
- (20) courtesy, pleasantness, friendliness, kindness, flattery.
- (21) laziness, over-eating, dissipation, extravagance, smoking.
- (22) thieving, cheating, forgery, unfairness, begging.
- (23) perjury, lying, swearing, slang, deceiving.
- (24) bullying, stealing, blackmail, cheating, swindling.
- (25) frankness, sincerity, simplicity, trustfulness, pliability.

Mental Survey Tests—Primer Scale

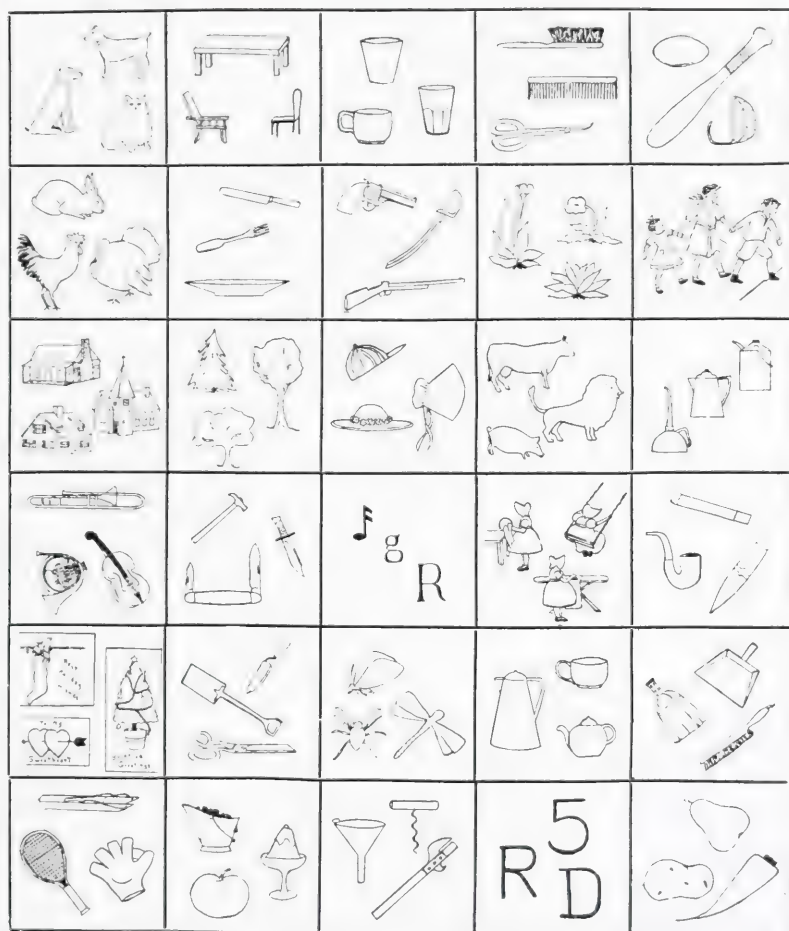
Schedule F

Name _____

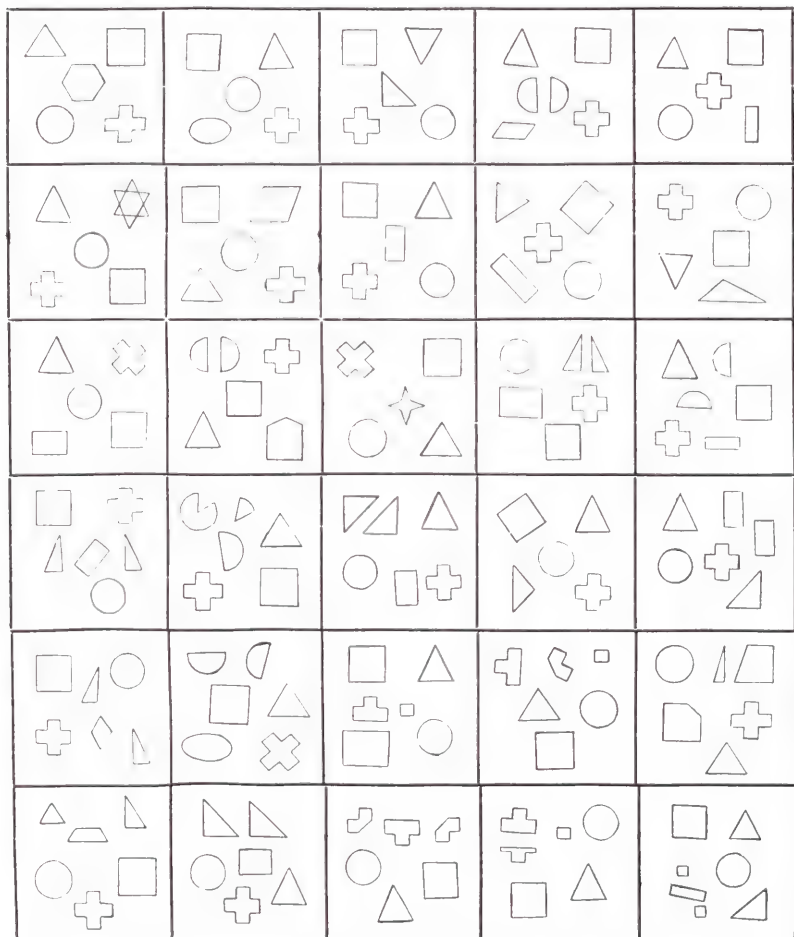
I. DOT PATTERN TEST



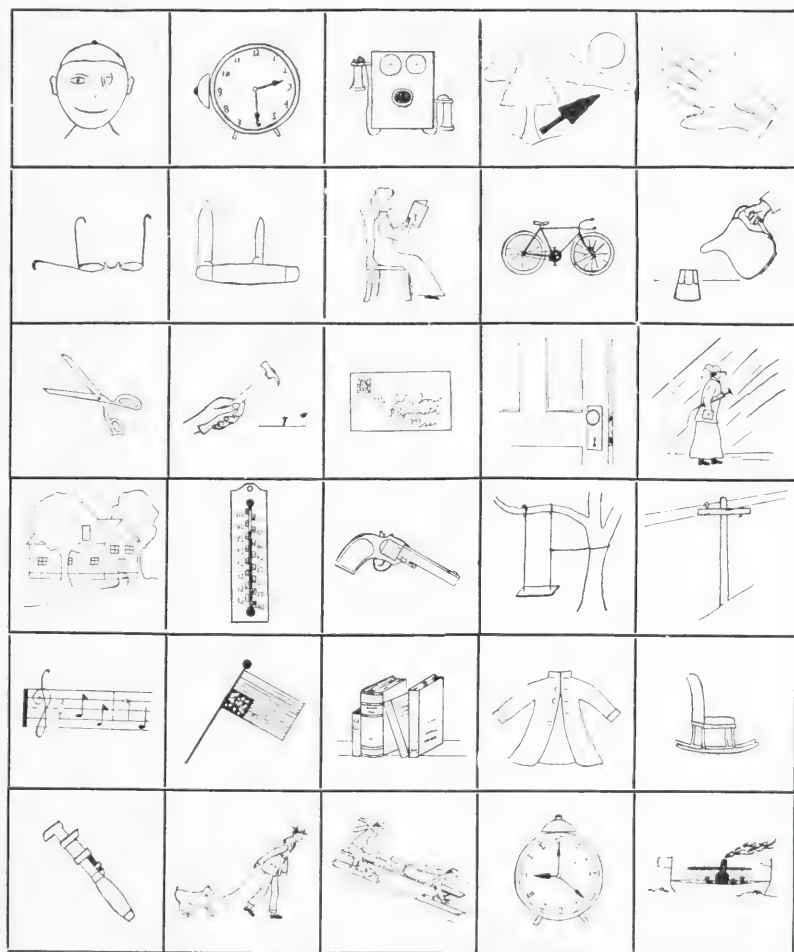
3. CLASSIFICATION TEST



III FORM BOARD TEST



IV. ABSURDITIES



DIRECTIONS TO EXAMINER

The examiner must go very slowly and make sure that each step is understood by all. However, there must be no lagging; the entire examination must not take more than twenty-five minutes, and should be given in less time in the third grade, where the directions can be read more rapidly. No one should attempt the examination until he has thoroly mastered the technique, as the slightest hesitation on the part of the examiner is immediately reflected by bewilderment and confusion on the part of the children. Every effort should be made to make the examination as informal, and as much like a class exercise, as possible. No questions should be permitted during the examination, however.

The children must be constantly watched for copying, as they are not yet old enough to understand that they should not do this. The examiner should always have on hand a supply of pencils; and no child should be allowed to lose time because a pencil breaks during the examination. If a child comes in late, leaves the room during the examination, or has his work otherwise interfered with, his paper should be distinctively marked at the time by the examiner. Note should also be made of any children troubled with deafness, poor eyesight, or any special defect.

It is well to have the teacher (if the examination is not to be given by her), *during* the examination, make out a rank order estimate of the ability of her pupils. Primary children are very dependent upon their teacher and, unless she is definitely removed from the examination, will look to her for permission every time the examiner gives a direction.

In giving the examination the examiner should have his own copy of the blank on which he should show the *general* position of the first example of each test, locating it in the upper left-hand corner of the page.

DIRECTIONS FOR SCORING

1. In each test the first five problems across the page are used as examples; do not count these in the score. In each test, each item correctly marked counts one point; the highest possible score in each test is thus 25.

2. Any mark (whether a single line, a cross, a dot, a line drawn around the object, or any other mark) that clearly indicates the correct object or part should be scored as correct. In general, when scoring is doubtful, give the child the benefit of the doubt.

3. If more than one dot in a group (Test I), object in a square (Tests II and III), or part in a picture (Test IV) is crossed out, score that problem 0.

The total score on the examination is the sum of the scores on the four tests. Arrange the papers for each room in order of total score and record the results on the opposite side of this page. Be sure to fill in all the information called for. Under age put age at last birthday; under grade put grade and half grade, and F for fast section, S for slow section, if the class is thus divided.

Tear the page in two on the dotted line; keep the "Teachers' Record" and return the other half to the Department of Psychology, Indiana University, Bloomington, Indiana.

DIRECTIONS FOR GIVING THE PRIMER SCALE¹

(Have the blanks already folded. Pass them out quickly, not giving the children time to study the tests or talk about them before the examination.)

Now is there any one of you who has *not* a blank and a good, sharp pencil? Hold up your hand.

Write (or print) your name on the line at the top of the page, then write your grade in the next space, and in the last space your age. (If the children cannot write, the teacher should be given the blanks some time before the examination, and should have them labeled and ready to give out to the children immediately at the time of the examination.)

Now attention! Look at the first page of your paper—the page that has the dots on it...And be sure you have the page so your name is at the top.

TEST I

Look at the first group of dots at the top of the page. (Hold up copy and indicate general position.) Can anyone see an extra dot—one that is out of place?...Where is it?...Yes, above all the others. Now I want you to take your pencils and cross out that extra dot. Just draw a line thru it. (Make sure that *all* the children get this first example correctly crossed out.)

Now look at the next group of dots, just *beside* the first group. Where is the extra dot?...Yes, below all the others. Everyone cross out the extra dot. Draw a line thru it.

Now look at the next group, just beside the last one. Where is the extra dot?...Yes, above the line. Now all cross it out.

Now look at the next group. Where is the extra dot?...Yes, the last one on this side (gesture to the children's right). Cross it out.

Now look at the last group. Where is the extra dot?...Yes, the one just below the middle dot. Cross it out.

Now, everyone attention! Finish the other groups of dots on the page in the same way. In each group there is just *one* extra dot. Cross it out. (Time allowed pupils, 3 minutes.)

Open the blank into the middle. Fold it over—so. Lay it down on your desks so the pictures are on top. (Do this with the children so that they may copy your motions.)

TEST II

Look at the first picture. What do you see?... (If the children do not name all the objects at once, keep asking, "and what else?" until everything in the square has been named.) Is there anyone who does

¹These directions are slightly modified from those used at the Conference, and are for the final revision of the scale, not for the form presented.

not see the two dogs and the cat?... Which two things are most alike?... Yes, the two dogs, and which one is different from the other two?... Yes, the cat. Now, I want you to cross out the cat, because it is different from the other two.

Look at the next picture. What do you see?... What two things are alike?... What one is different?... Yes, all cross out the cup.

What two things in the next picture are alike?... Which one is different?... Yes, cross out the curved (or crooked) line.

What two things in the next picture belong together?... Yes, the comb and brush. And what thing doesn't belong with them?... Yes, cross out the scissors.

Look at the last picture. In what game do you use the bat?... (If the children answer "ball game" ask "what kind of a ball game?") In what game do you use the mask?... In what game do you use the ball?... Then, what two things do you use in the same game?... Yes, the bat and the mask. And what thing do you use in a different game?... Yes, the ball. All cross out the ball.

Now, everyone attention! Finish the other pictures on the page in the same way. In each square, cross out the *one* thing that is different from both the other two. (Time allowed pupils, 3 minutes.)

This time do not open your blank, just turn it over to the page that has the little blocks on it.

TEST III

Look at the four places at the top of the page (point to the four forms). Do you all see them?...

Now look at the first square. Do you all see the little 3-sided piece in the upper corner?... Will it just fit into any one of the four places at the top of the page?... Which place?... Yes, the first one. Will the square piece beside it fit in?... Which place?... Will the cross?... Which place?... Will the round one?... Which place?... Now you have filled up all the places at the top, haven't you?... And what piece in the square have you left over?... Yes, the piece in the middle; it won't fit in anywhere. So cross it out.

Look at the next square. Will the first piece fit in?... Will the 3-sided one beside it?... Will the round one?... Will the cross?... And what piece have you left over?... Yes, the one that looks like an egg or a football. Cross it out.

Look at the next square. Will the first piece fit in?... Will the 3-sided piece fit in?... Will the round one fit in?... Will the cross fit in?... What one have you left out?... Yes, the one right in the middle. Cross it out.

Look at the next square. Will the top two pieces fit in?... What can you do to the two middle pieces to make them fit?... Yes, put them together. And where will they fit?... Yes, in the round place. Will the cross fit in?... And what one have you left?... Yes, the little flat piece, at the bottom; cross it out.

Look at the last square. Which piece won't fit in anywhere?... Yes the long, thin one in the corner at the top. Cross it out.

Now, everyone attention! Finish the other squares on the page in the same way. In each square, there will be *one* piece left over after fitting all the other pieces into the four places at the top of the page. Cross out the *one* piece that you can't use. (Time allowed children, 3 minutes.)

Turn over to the last page that has some more pictures on it.

TEST IV

Look at the first picture. What is wrong with the boy's face?... Yes, the eye. Cross out the eye, because it is wrong.

What is wrong with the table in the next picture?... Yes, the table has five legs. Which one is wrong?... Cross out the middle leg.

What is wrong with the telephone in the next picture?... Yes, there are two receivers. Which one is wrong?... Yes, the lower one. Cross it out.

What is wrong with the boy?... Yes, his foot is turned the wrong way. Cross out his foot.

Look at the last picture. Are you looking at the back of the glove or the front?... Yes, the back; and does a glove usually open in the front or in the back?... Then what is wrong?... Yes, the opening is on the back, so cross out the opening.

Now, everyone attention! Finish the other pictures on the page in the same way. In each picture cross out the *one* part that is wrong. (Time allowed pupils, 2 minutes.)

(Collect the blanks at once, not permitting any time for further work.)

(Wherever ?... is used it indicates that the children are to answer.)

A Group Scale of Intelligence for the First and Second Grades

MRS. SIDNEY L. PRESSEY, Graduate Student, Department of Philosophy,
Indiana University

IN the survey work which my husband and I have been doing this last year, we have been confronted with two main and very distinct problems. The first of these is to determine the kind of "pupil material" with which a school has to deal. This "pupil material" is sometimes of high native ability, sometimes of poor native ability; it is sometimes very homogeneous in its native endowment, and sometimes very heterogeneous. The first problem is, then, to determine the nature of this "pupil material". The second problem is to determine what the school has done with the material that has come to it.

Educational measurements have already made us familiar with the idea of group testing to discover to what level of attainment the school brings its material. We test progress in reading, writing, and geography, and in almost all other subjects, and we do it almost wholly by group testing. But there has been little systematic group measurement of the material with which the school has to deal. Yet I am sure that most of you would agree that schools may vary markedly in the ability of the children who attend them. In fact, in our survey work last year we found many striking differences between schools in the same system. For instance, one school tested over a year above another school which drew from a different neighborhood, but in the same system. It would seem then a logical step to measure first the material with which a school has to work before determining the methods it shall use or the results it shall obtain. Any fundamental differences in natural endowment should certainly be known to superintendents and teachers and should be taken into account in judging the accomplishment of these schools and in any complete understanding of their many problems.

But we found it exceedingly difficult to obtain satisfactory measures of these differences if tests, which could be given only in the upper grades, were used. Grade selection in any grade above the first profoundly influences the make-up of the grade. For instance, in one school recently tested there was twice as much retardation and only half as much acceleration as in another school in the same system. It is evident that the children in the upper grades of one of these schools were a far more highly selected group than the children in the upper grades of the other school. Therefore, a test based on the upper grades only is but a partial statement which is *quite* as dependent upon school selection as upon the natural abilities of the pupils. The natural way to

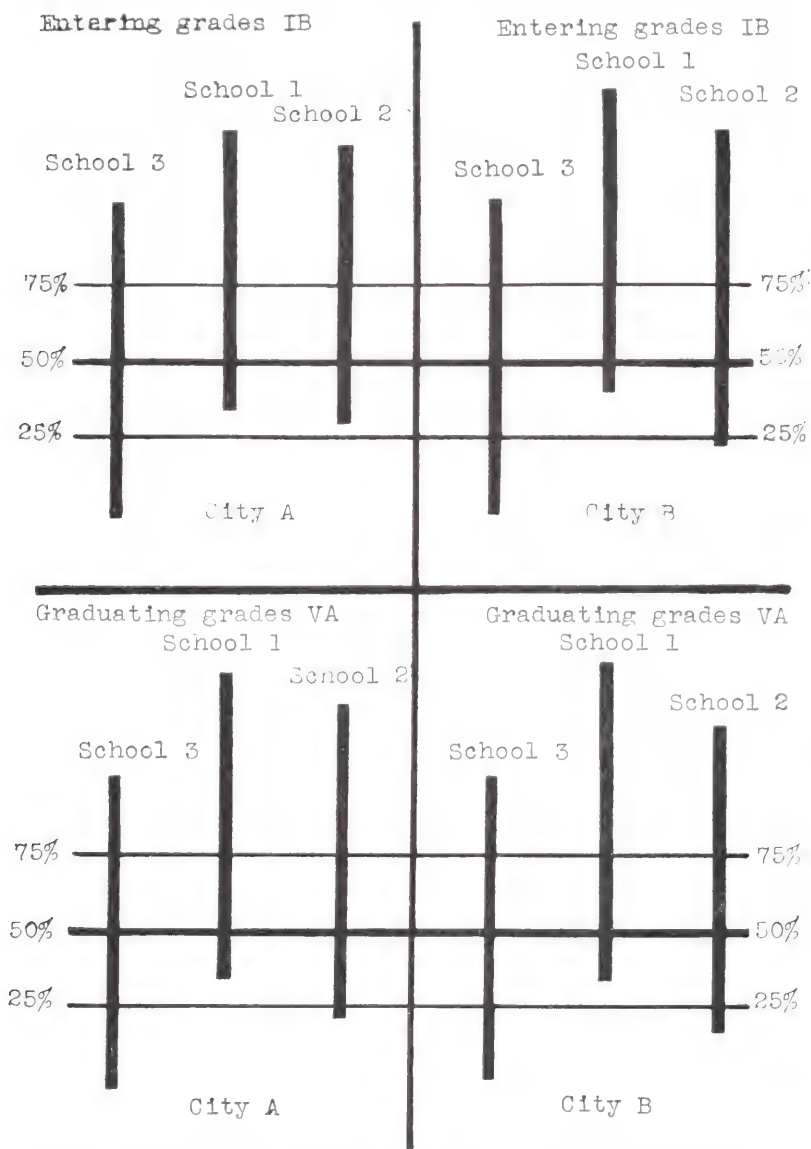
overcome these difficulties and obtain a true measure of the "stuff" with which the school begins its work was to devise a scale for measuring the abilities of the children at entrance, or soon enough after entrance so that these influences of selection and training have not had time to operate. The Primer Scale which I demonstrated a short time ago was the result of our attempt to solve this problem. I wish now to give you briefly certain facts which we have thus far found to show that these differences are really there and that the scale does measure them.

The results may be either in the form of group comparisons or individual diagnosis—the former being by far the more important and reliable of the two. Group comparisons may result in (1) comparison of entering classes—or of all the six-year-old children in a school—to determine the general level of ability; (2) a comparison of the range of abilities within a given grade or among these same six-year-old children; (3) a comparison of strong and weak sections within a single school with the idea of making each section a more compact unit for teaching purposes. As regards individual diagnosis, there are the usual two possibilities: first, the locating of bright and dull children for more intensive study; and, secondly, the locating of children who are misplaced and belong in some other grade than the one they are in. I shall take up these points in the order mentioned.

There are, in each of the two cities upon which these results are based, three ward schools. I shall be chiefly concerned with the first of these two cities, which I shall call City A. Two of its ward schools draw from the best sections of the city, but each one also draws from a neighborhood where the poorest families—the trash of the city—have their homes. The third ward school draws from a neighborhood of self-supporting, laboring people,—a neighborhood which has no homes of professional or successful business men, but which, on the other hand, has only a few destitute families. Since the first two schools I mentioned are somewhat similar in the make-up of their population I shall use the first of them only as a basis of comparison with the third school. In the other city, City B, I have selected the best and worst schools for comparison.

By both age and grade the third school (the poorest one) in City A, tests greatly below the median for the first school (See Figure 1). The difference in median is equal to the difference between the six-year and the seven-year medians, or just equal to a whole year in "mental age". This third school has, then, definitely inferior stock with which to deal; yet it is supposedly using the same methods and teaching the same subjects as the school whose median is more than twice as high as its own. The results of this policy are observable even as early as the first, second, and third grades (which were all tested by this scale). This third school, in these three grades, averages 26 per cent retardation,—twice as much as the first school, and has no acceleration whatever, while the first school averages 13 per cent acceleration. In other words, the poorest school is attempting by excessive retardation to bring its inferior pupils with the *same* methods to the *same* achievement as the

FIG. 1. A COMPARISON OF THE ENTERING (IB) AND GRADUATING (VA) GRADES IN THE THREE WARD SCHOOLS OF TWO CITIES.



The comparison in each city is with the poorest school (School 3). The other schools are shown as having so much in per cent above the median for the poorest school. The graduating grades of the three schools in each city will furnish pupils for the *same* grade of one junior high school.

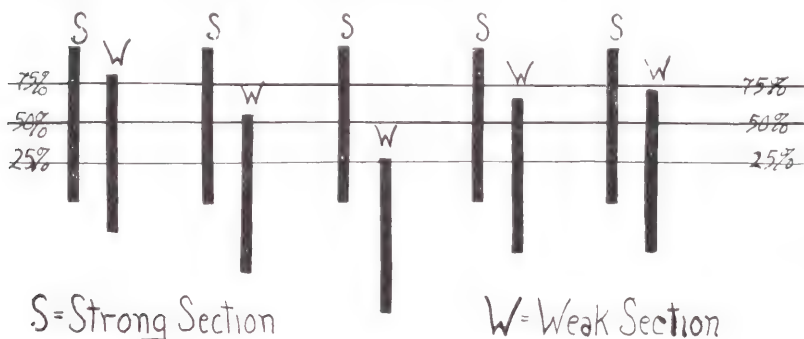
pupils in the best school. The fact that the grade norms are twice as far apart as the age norms would evidently indicate that this has not been done. This policy of excessive retardation, which is as high as 50 per cent in the fourth and fifth grades, is apparently the only attempt at adaptation to its inferior material that the school has made. Very vital questions are raised by such a situation. Should the goal of achievement be always the same? Should the methods be always the same? If the school's reaction in retarding its pupils is not sufficient to meet this difficulty, what else is there that the school can do? Whatever the answers to these questions may be, it is evident that the schools in which the pupils are of inferior stock—and there is usually at least one such school in every community—must find some way of more adequately adapting their curriculum and methods to the neighborhood which they serve. These two schools, then, in City A have very different problems as regards the general ability of their pupils. The same thing is true in City B. In this city 65 per cent of the primary children in the best school average above their age medians for the primary children in the worst school. The worst school averages, for its three primary grades, 40 per cent retardation, by means of which it has brought its grade medians very nearly up to the grade medians of the best school. It may be, then, that, while the fundamental situation is somewhat the same in City B, it is possible for them to make adequate adjustment to their pupil material by merely retarding their pupils. In City A the adjustment by mere retardation has very evidently not been sufficient. But in both cases, aside from any reaction on the part of the school, whether adequate or otherwise, the basic situation of a marked difference in the pupil material entering the first grade of the first and third schools remains unchanged. The difference in native ability is there, and the scale has measured it.

Turning from the general average of ability among the pupils entering different ward schools in the same system to the range of ability, as indicated by the tests, equally striking differences are found. The first school in City A, drawing from the very best *and* the very poorest elements of the city, has a range among its six-year-old pupils which is 146 per cent of the corresponding range in the third school, which draws from the laboring classes only. It is evident then that this third school in spite of its poorer median showing does not have some of the difficulties which confront the first school. There is an occasional very brilliant or really feeble-minded child, but the school is able to take care of them by acceleration and retardation. There are not enough of either class to be much of a problem. Their problem lies more in the generally inferior ability of *all* the children. The first school, however, finds an astounding range of ability among its six-year-old children. The school has already attempted to meet the situation. It shows for its primary grades 13 per cent acceleration and 13 per cent retardation. There has been then an effort to set off these two extreme classes; but the attempt has been only partially successful, as nearly 50 per cent of the entire population of the school belongs in either one or the other class. There is thus in this first school a clear demand for special classes

for its brilliant and subnormal children. The wide range of ability in the first school presents just as much of a problem as is presented by the generally inferior stock of the third school, and calls for just as thoro an adjustment on the part of the school, if the school is to serve its neighborhood to the fullest.

In City B a similar situation, tho a somewhat less striking one, is found. The six-year-old children in the best school show a range of ability which is 122 per cent of the range in the poorest school. The chief difference between the situation in the two cities lies in the lack of adjustment in City B. In its best school, with its considerably wider range, the acceleration of the primary grades amounts to only 7 per cent. For the fourth and fifth grades it is only 1 per cent. There is, however, 23 per cent retardation in the primary grades. The school has apparently made a more than necessary adjustment at the lower

FIG. 2. A COMPARISON OF THE STRONG AND WEAK SECTIONS IN THE PRIMARY GRADES IN A WARD SCHOOL IN CITY A.



The comparison is made on the basis of the strong section, the weak section being shown as so much in per cent above or below the median of the strong section.

end of its distribution, but has made no attempt to take account of the superior abilities of some of its pupils.

Wholly aside from the adjustment or lack of adjustment, or particular kind of adjustment, found in the two cities, the underlying problem is the same. It is a problem which would be found in the majority of cities, especially in small cities. A thoro adaptation on the part of any school to its pupil material must be preceded by accurate knowledge of the distribution of abilities among its pupils, as well as of the general level of their abilities. To gain this accurate knowledge some complete mental survey is necessary. Here again it becomes necessary to measure the "pupil material" before determining upon methods to be used or upon the results which are to be obtained.

Many of the primary grade rooms are divided into strong and weak sections. On this chart (Figure 2) I have indicated in diagrammatic form some of the strong and weak sections in the first grade rooms of

City A. All the sections could not be included, but these five are typical. For the most part, while the best section measures as distinctly better than the poorer section, there is considerable overlapping. The data presented in this chart show in the first place a real difference in ability, as measured by the scale, between the strong and weak sections. It also suggests certain shifting of small groups of children so as to make each section a more compact unit for teaching purposes. The scale should, I believe, be of distinct value in this connection, in pointing out the differences between sections and in helping the teacher to make a more accurate regrouping of her children according to their abilities. I am hoping to have under way very shortly an experiment in a certain ward school, in which in some of the rooms the children will be grouped in sections on the basis of the results from this scale—a Binet examination being given whenever the scale results and the teacher's estimate radically disagree. In other rooms of this same school the sections will be formed in the usual way on the basis of teachers' estimates alone. And at the end of the experiment I shall attempt to measure the comparative accuracy of the two methods of selection as far as achievement is concerned.

There remains one other aspect of the results from the Primer Scale to be discussed—their possible value for purposes of individual diagnosis, in the strict meaning of that term. The scale ratings give an average correlation with the teacher's ranking of her pupils in order of their ability, as she estimates it, of .58. The third grades were tested with both the "Cross-out" Scale and the Primer Scale; the correlation between the findings on the two scales was .64. When the scores of the pupils are grouped according to their ages it appears that a large majority of the accelerates rate in the uppermost 10 per cent of the distributions, and most of the retards in the lowest 10 per cent. There is thus some evidence to indicate a fair reliability of the scale in the individual case. But it must be remembered that the children are very young and are easily distracted by breaking a pencil, or by some other incidental happening during the examination, and that too much must not be expected of any examination under such circumstances.

It is especially interesting to notice, in studying the individual scores, that in all the schools, but especially in those schools where the retardation is highest, there appears an occasional retard who tests at the head, or nearly at the head, of his class. These children are usually found to be of the sort described by their teachers as "bright enough about some things" but unable to learn arithmetic, or reading, or spelling, or some other formal phase of school work. As you have probably noticed, the Primer Scale does not involve any formal knowledge on the part of the child. The examination is based instead upon observation of the happenings of everyday life. That certain children who are retarded in school do extremely well on such an examination raises some rather fundamental questions regarding the school curriculum. It may be that these children are essentially "misfits". But is it not quite as possible that the school is to blame? A curriculum based largely on the "three R's" may not be the curriculum which is best fitted to the special abil-

ities and aptitudes of these children. They are evidently slow in mastering the formal elements of language and numbers; but yet in matters of everyday experience they show a more than average amount of "common sense". It would seem not impossible that a better adaptation on the part of the school to the special needs of such cases might result in their rising from the group of chronic failures and doing *some* kind of work well—at least well enough to maintain their self-respect.

So much for the results we have obtained with the Primer Scale. In the brief moment which remains to me I wish to point out certain important general characteristics of the scale itself. It is composed of four tests which present, *without the use of either words or numbers* (that is, without requiring of the child any literacy whatever) a total of one hundred problems involving sense perception, general information, ideation, mechanical ingenuity, reasoning, and judgment. The scale takes about twenty minutes to give. When we tested Bloomington last week my husband and I began to test when school opened at 8:30. We continued examining until 2:30 (not testing, of course, during recess periods or at noon). When we stopped we had examined all the first-, second-, and third-grade children in Bloomington—a total of over six hundred cases. At 2:30, two of the students and I started scoring the papers. At 6:00 p.m. all the papers were scored. The complete survey was, then, made by two examiners and scored by three people in only one day. The convenience of the scale for the ready measurement of groups would seem evident.

There is one other feature of the test structure upon which I should like to comment. The four tests are all based upon the "Cross-out" principle—that is, in each test the child is asked to cross out an extra or wrong element in the situation. This method involves a minimum of labor on the part of the child. It also permits directions which are short and uniform thruout the test, thus avoiding the fatigue which small children show very quickly if they are shifted constantly from one problem to another, and if each test has a distinct and special set of directions. The naturalness of the problem from the child's point of view should also be noticed; he is asked to cross out something that is wrong or extra. When a child makes a wrong or extra mark on his paper he immediately crosses it out. One task of the primary school teacher is to make him abandon this method in favor of an eraser. On the mechanical side, the scale is thus distinguished by the naturalness of the recording act which it requires of the child; and with this naturalness there go also unusual ease in giving and readiness in scoring.

The paper may be briefly summarized. The Primer Scale is a group scale of intelligence, for use in the first, second, and third grades, but intended especially for the measurement of the raw "pupil material" as it first comes to the schools in the entering classes. The scale has thus far been given to a total of about 1,200 children—all the pupils in the first, second, and third grades of two small Indiana cities. It was found that the children entering the different ward schools within a given city may be markedly different as regards both general average of intelligence and range of ability. The children entering one school may test as high as a year in "mental age" above the children (of the

same chronological age) entering another school of the same city. Or, the range of ability in the entering class in one school may be half again as great as the similar range in another school of the same system. In the one case it is suggested that, if the second school is to be of the greatest possible service to the inferior pupils who attend it, there should be special adaptations, in curriculum and methods, to the problems presented by this inferiority. In the other case it is urged that the first school should make some special provision for its large number of a-typical (extremely brilliant *and* extremely dull) children. But the point I have tried to emphasize is this, that before a school can either clearly understand its problems or wisely meet them it must first obtain some measure of the "material" it has to work with—must determine the inherent possibilities of its children.

Two other uses of the test results are also suggested. In schools where the classes are divided into "strong" and "weak" sections the findings may assist the teachers in making these divisions more satisfactory. And the scale may help in individual diagnosis; these last two uses are, of course, related. But it is stated that while the reliability of the scale in the individual case would seem fair, the chief value of the scale should appear in the making of group comparisons rather than in individual diagnosis.

And let me urge for the last time that some measure of the "pupil material" with which a school works must be obtained before that school can either decide wisely upon methods or judge fairly of results. Once this knowledge of "pupil material" has been gained the school can begin, with intelligence and understanding, to make adjustments to its problems. It is with such adjustments that my husband's paper, which follows this, has to deal.

School Surveys by Means of Group Tests of Intelligence

SIDNEY L. PRESSEY, Research Assistant in Psychology, Indiana University

A YEAR ago we presented to you a somewhat elaborate group scale of intelligence, which we considered to be of a very considerable reliability in the individual case, and which we attempted to make particularly adaptable by special methods in giving the directions and by an unusual variety of problem and method for work with grade school children. You may be interested to know that the scale has aroused not a little interest and that it has been used extensively both in practical educational work and in research. This scale was used by us shortly after the Conference last year in some rather extensive school survey work here in the state, and certain results of these surveys were—to us at any rate—very surprising.

They suggested to us that there might be in educational work a distinct place for two types of group scale of intelligence. One type of scale should, like the scale we presented to you last spring, be organized primarily for the purpose of obtaining reliable ratings on the individual case—primarily for individual diagnosis. But there also seemed to be a distinct place for a brief group scale, in form especially designed for convenience and standardization in giving and readiness in scoring, to be used for *group* comparisons. We have already mentioned this at the demonstration, but I am anxious to bring the point up again, because this notion of the use of a scale of *intelligence* for group comparisons is, so far as I know, distinctly unique in educational work. We found in these group comparisons last spring distinct differences between the children in different neighborhoods or communities, in mental endowment. Still more striking were the mass differences between different schools and school systems in the ways they dealt with the problems presented by these differences in ability among their school children. The results have been summarized briefly in a recent article of mine in *Educational Administration and Supervision* which some of you have doubtless seen, and need not be detailed here.¹

They gave us our cue for our work this year. We have attempted, this year, to develop two brief group scales designed primarily for the measurement of groups. One of these scales has been planned for use with first- and second-grade children—that is, for use with the children

¹"A Comparison of Two Cities and Their School Systems by Means of a Group Scale of Intelligence." *Educational Administration and Supervision*, Vol. V, No. 2, February, 1919.

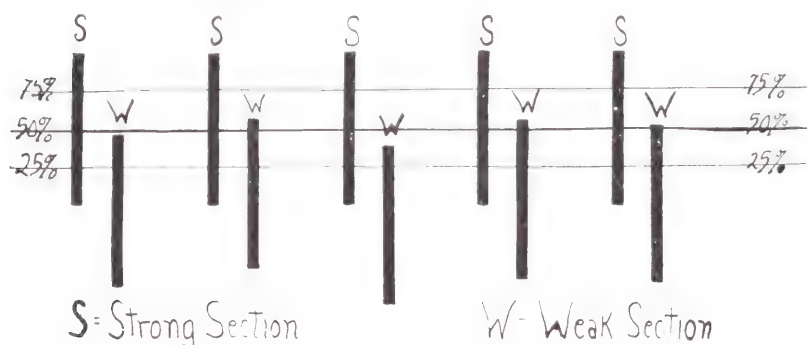
as they enter the school. Our results along this line have just been presented to you by my wife. Her scale has been designed primarily to give a measure of differences in "pupil material" as that material comes to the school. The scale I gave you at the beginning of the afternoon was designed primarily for a different purpose—for the measurement of the adaptation which the school shows, in the later grades, to the problems presented by its "pupil material" and the individual differences which it finds among its children. And I wish briefly to illustrate to you three ways in which it seems to me such a scale may be of value in educational work. I shall present to you three groups of facts. The first will indicate something of the adaptation found in a single school, as measured by these tests. The second group of facts will deal with problems of inter-adjustment in different schools in the same system. And the third set of facts will have to do with differences between different school systems in their methods of adjustment. Most, tho not all, of these results have been obtained with the "Cross-out" Tests. But that does not matter now. The question is as to what uses *any* such scale of intelligence might serve in educational investigation.

First as to evidences which have been obtained, by the use of such a group scale, regarding the adequacy of adjustment to individual differences in mental endowment within a given school. To put the matter in another way, what might a principal hope to obtain by a survey of his school with such tests? Such a principal might find his scores most useful if he should tabulate them taking account of age-grade status. His median scores at each grade are much less important; they are determined by factors largely not under his control, like the average mental endowment of the children in the neighborhood from which his school draws. But it is important for him to know whether he selects his children accurately for whatever standard he may have, and whether the children in each class are sufficiently alike in ability to make a convenient teaching unit. I am sure that you will all agree with me that excessive retardation is unfortunate, and that no child should be kept in a grade below that in which he is capable of accomplishing the work required. We have put the matter something like this: suppose a child is retarded but tests above the median for the grade he is in. Would it not be well to stretch a point in such a case and see whether he might not do passably in a grade above the one he is now in, among children more his fellows? Not that the ratings obtained with this scale need be considered sufficient evidence for advancing that particular boy, you understand. But by the number of retards testing above the median for the grade in which they are placed we should obtain some measure of the extent to which retardation is consciously worked against, and the problem of adequate placement satisfactorily solved. It is hardly possible to present the facts in the graphic form I have tried to make use of in showing our other data. But I may say shortly that School A (as we will call it) shows 30 retards above the medians for their grades, and another school in the same system shows 15. School A shows 5 retards above the median for the grade next above theirs; School B shows only 1 such case. May I suggest that there would ap-

pear to be an excessive amount of retardation in School A (there is just that) and that the retardation has been somewhat indiscriminate? You may also be interested to know that the grade distributions for School A are 28 per cent wider than for School B. That is, School A not only retards more, but selects less accurately, as well, for its grades. May I suggest that definite action on the part of the principal of School A to make his grade adjustments more accurate is—at least—suggested.

But such matters of adjusting school work to the capacity of a child come to explicit recognition most fully in junior high school. And here I want to present very briefly indeed and schematically our results with the "Cross-out" Tests from the junior high school of another small Indiana city. This chart (Figure 1) shows girls' sections for grades from 6 to 8 inclusive—each half-grade except one having the girls divided into two sections on the basis of ability as estimated by the teachers.

FIG. 1. A COMPARISON OF THE STRONG AND WEAK SECTIONS IN A JUNIOR HIGH SCHOOL (GRADES 6B-8A).



The comparison is made on the basis of the strong section, the weak section being shown as so much in per cent above or below the median of the corresponding strong section.

I wish you would notice in the first place that the groups are distinct thruout. This is, the tests evidently measure, to a very considerable extent at least, that sort of mental ability which makes a child bright enough for a good section or so inadequate in school work that he must be put into a slow section. But there is very considerable overlapping. That overlapping is most considerable in the 6A and 7A grades. Would it not seem plausible to use such a scale as a means of measuring the satisfactoriness of such division? It is not that the scale would necessarily indicate *which* children should be transferred (altho personally I am inclined to believe it is sufficiently reliable to aid in such a selection). But it would certainly give a means enabling a principal to compare sectional divisions, and to diagnose weak points in the total system—means of readjustment might be suggested by other methods.

I thought of calling this paper "The Group Scale of Intelligence in Educational Supervision". That it might find some service in super-

vision I have tried to indicate by this second chart (in lower half of Figure 1 of the preceding paper). City A and City B have junior high schools for the grades from 6 to 8. Each city has three ward schools. You see here indicated by these black lines the distribution of abilities in the fifth grades (the classes about to enter junior high school) in these three ward schools. You will see that in this first system the fifth-grade class in the poorest school shows only about 35 per cent above the median for the graduating class in the best school. In City B the worst school has only 30 per cent above the median for the same graduating class. Cities C and D have no junior high school, so we have compared eighth grades—the classes which will, in part at least, enter high school next year. We would expect that by the eighth grade there might have been sufficient time for the extra elimination and retardation usually found in poorer schools to operate sufficiently to make the graduating classes more alike. As a matter of fact we find our greatest difference here—only 20 per cent in the worst school above the median of the best school.²

Such differences, particularly in ward schools contributing to junior high school, are surely of sufficient importance to require study. And I should like to suggest that further work of this sort should be of not a little interest. You see here in this last city there is an example of a different problem: there is not only difference in average ability but difference in range of ability between the two schools. Such a difference would suggest a greater variety in the curriculum, or a greater flexibility in the organization, to take care of these differences. We have, as a matter of fact, made such school comparisons for entire schools from the second grade up. The farther down we go the more the differences are, in most instances. I have chosen the graduating classes because there, it seemed to me, the practical problem of adjustment became most acute, and the need of such measures most pressing.

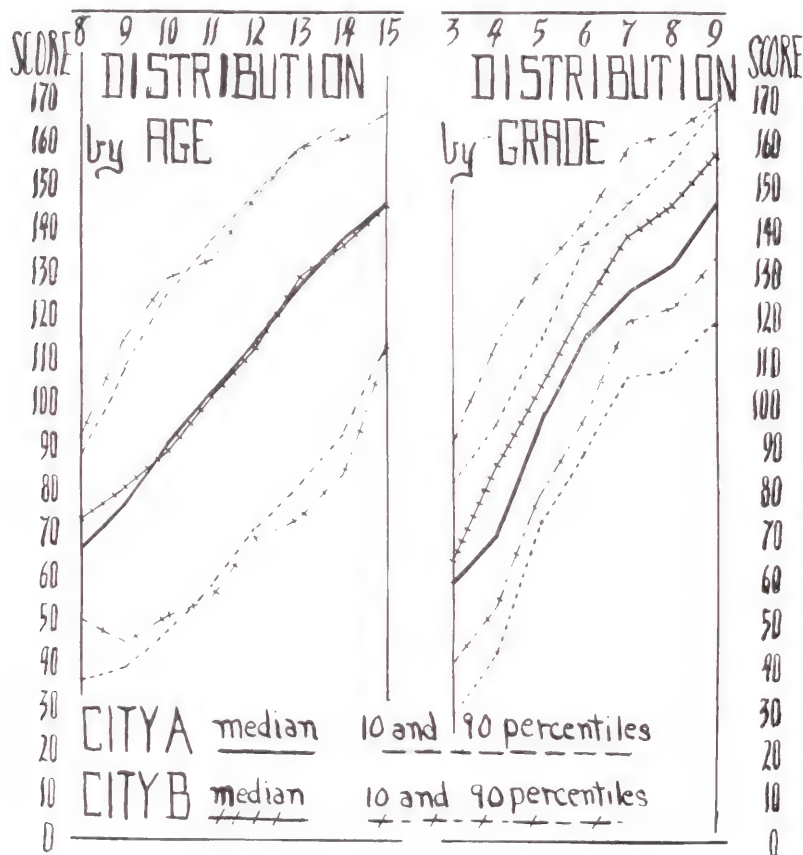
I wish I could give you more data about these school differences. They are intensely interesting, and reflect, I believe, interesting sociological as well as educational conditions. But I am very anxious to describe also certain important features of the two-city comparison—an outline of which was published in the article I referred to a short time ago. To present the facts most simply we have prepared these two charts, with a mimeograph curve, which some of you have already obtained (Figure 2). You see that age for age the children of City A and City B are, in the mass, almost absolutely similar. They average alike, and there are about the same number of extreme cases. But grade for grade, the distributions are strikingly different. As an actual fact altho the children were of equal ability, the sixth and eighth grades in City B averaged above the seventh and ninth grades in City A! The explanation is complex, but the important causative factors were undoubtedly to be found in a retardation twice as great in the second city, a quarter as much acceleration—and the excessive elimination that usually goes along with such conditions. The children of the second city averaged grade for grade over five months older than in the first

²Results from Cities C and D (shown on the original chart) are omitted.

city. And there were over a third less children in the seventh and eighth grade in proportion to the third and fifth grades, than was the case in City A.

The facts are surely striking. They suggest, it seems to me, something of a new approach to the old problems of retardation and grade adjustment. Not that a brief survey scale such as we have shown you

FIG. 2. A COMPARISON, BY AGE AND BY GRADE, OF THE ABILITIES OF THE SCHOOL CHILDREN IN TWO CITIES.



will suggest individual readjustments to remedy such conditions (altho personally I believe it can be of great value in pointing out the children who had best be considered in making the adjustment). But such extensive surveys give us a large perspective on the situation which enables us to deal with such problems much more wisely than would be possible otherwise.

To summarize then. The contention is, that there is a distinct service to be performed in educational work by a brief group scale of intelligence designed especially for comparing sections, classes, schools, systems or other *groups* with regard to the average and distribution of ability in each group. We have found some evidence for supposing that such tests might be of use to a principal in making his grade and section divisions more accurate. They should assist a superintendent in standardizing conditions in the different schools of his system. They should aid both superintendent and principal by making possible comparisons between systems and cities, thus giving a larger perspective by which to judge local problems. Will you permit me now to say just a word about the particular brief scale with which we have experimented, as useful for such comparisons?

I must first emphasize again (we are so used to thinking of tests of general intelligence as tests useful only in individual diagnosis) that this brief scale of "Cross-out" Tests is designed for the making of *group* comparisons,—exactly as are the majority of our tests of achievement. For a scale to be used in such extensive testing as I have suggested, the utmost possible in economy of time and effort in giving and scoring, and in compactness and convenience of materials, is a first essential. Otherwise, such surveys are altogether impracticable. There must be, further, rigid standardization of method and absolute objectivity in scoring, since the tests will be given in most instances by the teachers themselves and often scored by them.

First, as regards ease in giving and scoring. The examination never takes over twenty-five minutes to give. The directions are sufficiently short to be all placed on one side of a typewritten sheet; they are not elaborate, and are to be read verbatim from this sheet (I may say that in our work thus far the examinations have been given largely by the teachers). Scoring is absolutely objective; either the right word or number is crossed out or it is not. There are no half scores, no ambiguous scores, no alternative scores. The scoring is so easy that it is possible to score an entire examination in one minute. That is, a teacher should be able to give the examination to an average-sized class, and score the results all within an hour. I wish to stress this convenience. There are few achievement tests which can be handled so readily. In their studied adaptation to your convenience, I believe these two scales which we have presented to you today are somewhat unique.

You will wonder if an examination so brief can be adequate for the purpose in hand. I wish to point out in this connection the extreme condensation of matter and problem we have introduced. An examination of one hundred items is presented on two sides of a 9 by 12 inch sheet. And the problems are absolutely clear cut; the possibility of chance successes is negligible. I am very anxious to invite comparisons on these points. By way of contrast in this matter I have placed on the table by the door sample copies of the army scale Alpha, of a group scale of intelligence recently devised by Arthur S. Otis, and of a modification of the army scale now being used by Professor Haggerty in the Virginia Survey. I think you will see at once how we have gained in compact-

ness of form. The Virginia Commission presents practically the same number of problems on over three times as much paper; and there is involved, I should guess, at least twice as much labor in scoring.

Particularly have we obtained "condensation of problem". The first of the "Cross-out" Tests—the test of language ability—will serve as an example. Many of you have had experience with the Binet scale. You then know the form of "disarranged sentences" which appears in that scale. The child is shown three sentences with the words mixed up, and is asked to reconstruct the sentence orally. This is surely the most naïve form possible for such a test of language ability. The form requires individual testing. Timing is usually inexact, and directions loose. Finally there are no gradations in the score; the subject either passes the test or does not. An attempt to put the test in group form appears in the army scale Alpha. Here the subject is given a series of relatively simple disarranged sentences. Following each sentence are the words "true" and "false". The subject is told to figure out the sentence and then draw a line under "true" if the sentence is true or a line under "false" if the statement made by the sentence is false. For instance:

write are with to pencils.....	true	false
are and apples long thin.....	true	false

You will doubtless see at once that, in the first place, the matter which decides the subject's answer—the truth or falsity of the statement—is a matter beyond the understanding of the sentence itself. That is, we have a confusion of two things in the test,—ability to work word puzzles and judgment about certain common sense statements of fact. But altogether aside from this is the point that if you just guessed, you would get about half right by chance without ever solving the problem at all,—that is, half the blank is wasted.

Trabue has a different form of language test, with which you are undoubtedly familiar. He has ten sentences with certain words left out, and he allows the subjects seven minutes in which to fill in these missing words. After the words are all filled in, you know, if you have ever tried to score the Trabue completion scales, that a still worse task awaits the examiner. In this first test of the "Cross-outs" there are not ten but twenty-five sentences; yet the test is given in five minutes. It can be scored in less than fifteen seconds. And I am inclined to think that it will give, if anything, a more valuable measure of language ability.

We have tried to select these four tests as perhaps measuring some special ability of this sort besides being contributory to a total rating. The second might best be called a test of "information with understanding". I want you to notice that in each line there is a problem in four-fold similarity and fourfold difference. Those of you who are familiar with the Binet will again appreciate the great concentration of problem and condensation of matter as compared with the differences and similarities of that scale. About the remaining two tests I need say less. In the third we have hoped to find a test of potential arithmetical ability. The fourth is primarily a test of vocabulary and, in agreement with

Professor Terman's statement about his vocabulary test, we find it the most differential of the four.

This great condensation of problem is the most important single feature of the scale. Because of this condensation, and because of the very careful selection of tests which previous research has shown to be most highly differential of degrees of "general ability", we are able to present you with a group scale of "general intelligence" sufficiently convenient, and sufficiently standard in directions and objective in scoring, to make practically possible such extensive comparisons as I have suggested in the first part of my paper. The scale has now been given, in one form or another, to about four thousand cases. The items have twice been regraded and rearranged in the order of difficulty, and the form revised. The scale is by no means developed to the point I should like. But it will, I believe, serve very well the purpose I have had in mind—will prove an unusually convenient and ready method for first investigation of such problems of group comparison in the schools as I have outlined to you.

The question as to the particular merits of this particular scale of ours is, however, a comparatively minor matter. The matter I wish to stress is the importance of the problem for the further investigation of which we have planned these tests and the distinction between the use of tests of intelligence which I have suggested to you and the use to which you are accustomed. You are accustomed to the use of scales of intelligence for purposes of individual diagnosis. You have probably been amused at the somewhat ostentatious way in which I have avoided discussion as to the value of these group tests for that purpose. I believe they *can* be of use in helping to indicate the children who are misplaced in school. If John, who is a retard, does remarkably well on these tests that would surely suggest a reopening of his case, to discover whether somehow the teacher cannot reach him and obtain better work from him in the future. If Mary, who is in a slow section in junior high school, makes the best score in her section on these tests, that would suggest consideration of the question as to whether Mary might not be transferred to another section of her class. But with these matters of individual readjustment I am not concerned this afternoon. Instead I have tried to make the one point that if we are to understand the problems presented to the schools by the individual differences found among their pupils, or satisfactorily estimate the adjustments of the schools to these differences, we must first obtain a mass measurement of the amount of these differences and the extent of the adjustments. I believe there is a whole new field for investigation here, which has hardly as yet been opened. And I wish, if I can, to leave you with some sense of the importance of such work, and an interest in further investigations in this field.

Group Tests for Intelligence in the Bedford (Indiana) Schools

EMERY W. MONTGOMERY, Superintendent of Schools, Bedford, Indiana

AN Age-Grade Study was made in the Bedford schools in the fall of 1917 which showed considerable retardation and a very small amount of acceleration. Special coaching classes were organized to help remedy this condition, and the Binet-Simon Tests were used to select the pupils for these classes. We were anxious, however, for some adequate intelligence test that would answer the purpose and that would be much more practical. Therefore we were very glad to have Dr. and Mrs. Pressey come to us in the spring of 1918 and test 1,009 pupils ranging from the third to the ninth grade and a few selected from classes above that grade. This last winter they made another test of all our pupils, using one intelligence test from the fourth to the twelfth grade, inclusive, and a special test for the primary grades.

These tests showed a number of interesting facts. The tests of 1918 showed 75 children who were retarded in their age-grade standing but who made a rating in the intelligence test above the median for their ages. We also discovered 163 pupils from grades three to nine who, although retarded in school, gave intelligence ratings above the medians for their grades. Fifty-two retards showed a group test score above the grade in which they were actually placed. It was immediately seen of course that this was the cause of some of our elimination.

Another interesting situation was discovered. It was found that the children of one of the schools (School B we shall call it) averaged over a year in mental age below the children of School A. The retardation, however, was found to be only 8 per cent greater in the poorer school and the acceleration only about 5 per cent less. The better school had evidently not profited to the fullest extent from its higher mental age. The following figures will show that these group tests may be used to compare two schools as well as to compare individual pupils: in one school 216 children were tested and in the other 214 (from Grades 3 to 5 inclusive in both schools). School A draws most of its pupils from the best end of town. Forty-one per cent of the children in School A scored above the median for their school age, 8 per cent rated above the 90 percentile, and 12 per cent below the 10 percentile. School B draws most of its pupils from the poorer end of town. This school shows only 27 per cent of its pupils scoring above the age medians, 3 per cent scoring above the 90 percentile for their age, and 22 per cent scoring below the 10 percentile for their age.

The results of the tests given this year as between buildings are similar to the results of last year. School A again scored above the

medians for all grades, while School B continued to score below the medians for all grades. We find, however, that there are more pupils in the normal age group thruout the school system, due to a conscious effort to lessen retardation.

We have made important uses of the results of this test. We have used the test as an aid in classifying pupils. When pupils were found who gave intelligence ratings above the medians for their grades we have made an effort to place them where they belong and thus lessen our retardation, increase our acceleration, and consequently reduce our elimination. We have not only done this in isolated cases, but Miss Dorothy Moran, principal of the junior high school, has used the test as a basis for the organization of opportunity classes for purposes of acceleration. Such a group was recently promoted to the high school and the pupils are doing excellent work. Coaching classes have been established and cadets are used to assist those who are seriously retarded. Use has also been made of the tests to determine accelerated and retarded groups in the grades and to assist in the introduction of supervised study in several rooms.

Mr. M. J. Abbett, principal of the high school, has made use of these group intelligence tests along with the Free Association, Analogy, and Trabue Language Tests of Columbia University to establish accelerated groups in algebra, language, and English classes in the high school. He has also used them to select pupils for accelerated work in the commercial department and to recommend that boys take shop work in some cases. He has made further use of them to help determine whether the pupil has language ability and whether he should drop or continue his language when failing. They have even been an aid in discipline by determining whether a pupil is doing all that he is capable of. I may say here that we have found a perfect parallel in the results obtained from Dr. Pressey's tests and those of Drs. Trabue and Briggs, of Columbia University.

We desire to thank Dr. Book and Dr. and Mrs. Pressey for their great assistance in this work, and I wish to state that much of the material discussed here is taken from Dr. Pressey's report which he sent to me after scoring and working over the tests. We are continuing our work on the results of the tests and we intend to make still greater use of them in the future.

Mental Defectives in the Rural Schools of County H in Relation to Their Physical Environment

HAZEL HANSFORD, Research Fellow in Psychology and Sociology, Indiana University

IN the years 1917 and 1918, the Indiana Committee on Mental Defectives made a survey of ten counties for the purpose of obtaining accurate information concerning mental defectives in Indiana for the state legislature. County H was made the subject of special study and it is concerning the schools of this county that I wish to speak.

County H is situated south of the central part of the state in a wooded hilly region. The central part is composed for the most part of good farming land. The outskirts are rough and hilly, unfit for the general farming of the only kind done by the average farmer. Besides being rough, these districts are, for the most part, isolated from the more civilized portions of the county. It is in these isolated rural schools tucked in among the hills and valleys that the largest per cent of feeble-minded and backward children were found. It was expected that in certain districts the number of defectives would be higher than in others but the amount of difference was not suspected. In one township it was found that 27 per cent of the school children were feeble-minded. This is probably a conservative number as the testing was done at a time when many of the children were not in school and it is a well-known fact that the dull children are the first to drop out of school.

When the results were studied according to schools it was found that the high per cent of this township, which we shall call Number 1, was due to the large number of defectives found in two schools in the township, and that those schools were located in two of the most backward and unprogressive communities in the whole county. There is not time to go into the social and economic history of each of the school districts which I wish to mention, but I will take time to describe briefly the conditions of School District A so that you may have a picture of what I have termed an "unprogressive community".

District A of Township 1 consists of a church, two stores, and perhaps a half-dozen houses scattered along the road. The schoolhouse is one mile north of the settlement. To reach this village from any direction, it is necessary to climb a long, steep, slippery, gullied, clay hill. The surrounding hills are covered with forests but this one is barren and bleak. The people try to farm but nothing seems to flourish. The whole place gives the impression of dilapidation, sordidness, and miserableness. Yet the inhabitants are not really miserable—they are stolidly

indifferent to their lack of comfort and beauty. All the houses are log cabins or small box shacks. Some of them have partially fallen down. In others only one room is floored. They are all miserably cold in winter and have no protection from the heat in summer. In a number of the homes in this district, the field worker found chickens in the house with the family. One woman in particular is remembered as being seemingly unaware that it was an unusual thing for a hen to lay her egg on the family bed. The only comment made by her was that the hen laid in the same place every day. Dirty, listless children were everywhere. They did not seem to care to play either at home or at school. The women moved about in the same "don't care" fashion as did the children. They spent most of their time going from house to house with the latest gossip, never allowing their household duties to interfere with their leisure. The only clean house on the hill belonged to a widow who has spent most of her life elsewhere.

The schoolhouse was a dilapidated old building, one corner of which had been propped up to keep it from toppling. The school was taught by an old man said to have had no license who was filling out the unexpired term of a young lady who had given it up as a bad job. He was teaching the younger children to count to a hundred and to repeat the alphabet. After six months' practice some of the pupils were quite proficient in this task. In this class were a few children who had been in school several years without being promoted and these boys seemed to know their lessons no better than did those who were spending their first year in school. This old teacher spent no little time of the two days informing the field worker and the pupils of how he had "thrashed" boys worse than the present school held and how he could do it again.

The children present were tested by the Stanford Revised Scale. Out of the 19 children, 15 were feeble-minded. This at first thought seems incredible. The remaining 4 were average children, but compared with the rest of the school they were shining lights. Two of them belonged to the widow mentioned before as being the one good housekeeper in the settlement. There was a total enrollment of 37 pupils but the attendance was very irregular. According to the teacher, some of the pupils had been out all year. It seems that the attendance office had not yet penetrated Township 1. Not only were there a large number of children found to be feeble-minded, but there were none found who were more than average in intelligence. One child had an intelligence quotient of 96 and the others were all in the seventies. Where stock is uniformly inferior as in this place, it looks like an up-hill job for any school teacher.

Only one other school in this township is anywhere nearly so bad as School A and that is School B. In B, out of 12 tested, 6 are feeble-minded. The other 6 are not even average, 5 of them having intelligence quotients in seventy and one in eighty. There was an enrollment of 25, but, as in the other school, attendance was irregular. The general tone of this school was higher than that of School A as the children were all fairly clean and the dull, apathetic look common to the first school was not present in School B.

Another school where there is a large number of feeble-minded and backward children is School X of Township 2. This school is in a small, rough section of the township which was early settled by the common ancestor of almost every family now living in the community. There were 14 children in School X and of that number, 11 were tested. Six of these were feeble-minded and 4 were borderline.

In School Y of this same township is to be found another large group of mental defectives. Twelve out of 18 were tested and 6 of that 12 were decidedly feeble-minded.

In School M of Township 5 are to be found 5 defectives out of a total of 22 pupils. This school like the others mentioned is located in a rough, unfertile district where ambitious folks would not care to live.

There is another group of schools which for lack of time cannot be discussed here where the per cent of defectives is very small. These are all located in the midst of fertile farming country. In these districts practically all of the defective children are the offspring of farm tenants showing bad family histories. In Township 3, only 1 feeble-minded child was found in the schools and he had strayed in from a degenerate group in a neighboring township.

In looking for the cause of the large number of mental defectives in some of the schools it was found that often not only were the children closely related, thus receiving an inferior mental inheritance from the same source, but that many of the defectives in one school were related to the defectives in the adjoining school district. For example, the Johnson settlement is midway between Schools A and B of Township 1, School M of Township 5, and School R of Township 4. In each of these schools are to be found descendants of "old man" Johnson who originally settled there. Three of them who attend School A, 3 in B, 3 in School M, 2 in School R of Township 4, and 2 in School E of Township 9 are feeble-minded. Besides these there are a number of others who cannot be said to be feeble-minded but who have intelligence quotients somewhere in seventy. Among those who are feeble-minded, there is 1 intelligence quotient of 42, 5 somewhere in fifty, and 4 in sixty. Thus among the young descendants of the Johnson family are 11 defectives of whom we know besides a large number who probably carry an inferior heredity which they will transmit to the next generation.

The 6 feeble-minded children in School X of Township 2 are cousins related to a large group of epileptic, insane, and feeble-minded persons all of whom are the descendants of a John Allen. He came into County H in the early days and settled on Ripple Creek among the hills while his neighbors, whose descendants are now among the leading citizens of the county, took up the fertile farming land. Here, as in other cases, the inferior stock has not only been pushed into the hills early in the history of the county, but it has been kept there by the more complex environment on the outside. In this little community, this one family has intermarried with two or three other families who have also been pushed back out of the fertile farming lands until at the present time practically all the residents of this bit of rough land are related by blood ties and we find their children making up the list of mental defectives in the district school.

If there were time it would be possible to take every unprogressive isolated district and show that there has been practically the same population history as shown in the above brief family sketches. An Indiana University study which deals more fully with the mental defectives of County H and their problems will be published this year. In it, evidence has been produced to indicate that at least three things have had something to do with the great amount of mental defect in the rough rural district. (1) It is probably true that in the districts described the stock has always been inferior and in many places this is an established fact. The more enterprising settlers took up the best land and the inferior men who did manage to procure good farms were soon pushed back into the hills where competition was not so keen. Just how this was accomplished is shown in the study of the family histories of not only degenerate groups but also by the study of the histories of well-to-do families and of the part played by each in the development of the community life. (2) It is well known that individuals tend to marry in their own class, and as the years have passed the inferior families pushed back among the hills have intermarried until at the present time many of the residents of any one of these small communities can trace their ancestry to the same man. Thus, there is a bringing together of bad traits in the present generation which were serious enough many years ago to set the possessors of them apart from their neighbors. (3) There has for many years been a constant draining off of the most ambitious blood from the country to the city where there are to be had greater advantages. Consequently, the more inferior stock left behind tends to inbreed, producing more inferior individuals out of which number the best in their turn tend to move to the cities.

Some Experiments with Mental Tests as an Aid in the Selection and Placement of Clerical Workers in a Large Factory

CHESTER S. CARNY, Consultant with Scovell, Wellington, and Company,
Industrial Engineers, Boston, Mass.

Introduction. Three types of tests have been used in the selection of applicants: trade tests, general intelligence tests, and tests for special abilities.

Trade tests as used by the army are so well known as to require no description. They aim to discover whether an applicant is skilled, semi-skilled, or merely a beginner in any particular trade.

General intelligence tests are measures of certain general abilities which are supposed to function to some considerable extent in many lines of work. They are often called tests of memory, attention, observation, etc. Their use is based upon the assumption that there is a "general intelligence" which will aid its fortunate possessor to succeed in almost any line of work he may undertake. Given this magic endowment he may hope to become doctor or lawyer, banker or engineer, dependent only upon choice, chance, and early training. How far this assumption is true it is not easy to investigate. Many pages have been covered with arguments *pro* and *con*. It is like trying to prove whether the patient died because of the doctor's medicine or in spite of it.

Tests for special ability assume that there are some particular qualities which aid the salesman to sell, other qualities which enable the lawyer to win his cases, still others which help the architect to design bridges; that while there may be such a thing as general intelligence which would aid a man in any of these professions, yet the man who climbs up out of mediocrity does so not by an excess of general intelligence, but by virtue of certain special aptitudes which function in his particular profession.

The trade test answers the question, Is he a skilled mechanic? The general intelligence tests asks, Has he the intelligence usually possessed by skilled mechanics? The special ability tests seek to discover whether the applicant is endowed with those natural gifts without which he can scarcely hope to become a first-class mechanic. Trade tests can be used only on men who have worked as mechanics at the trade. The other tests can be applied to men who have never tried that kind of work. How much money is being wasted in trying to make surgeons out of men who should be blacksmiths, or accountants out of men who should be carpenters, who can estimate? Certainly it is high time the employer and the educator attacked the problem of placement and vocational guidance with all the tools that science has made available.

Two Methods of Attack. For the investigator who wishes to obtain a set of tests for any particular task, two points of departure are possible. He may start with the simple mental tests, covering as wide a range as possible, in the hope that some few in the set may by good fortune show correlation with the required task. Scott's tests for salesmen were built in this way. The process is often long drawn out, and disappointments are many.

The second point of departure is an analysis of the task itself. Each operation is reduced to its simplest factors, and these are standardized into tests. Such a sampling process was used by Münsterberg in his tests for motormen and ship captains. It was also followed in the army tests for truck-drivers.

By whichever method the tests are obtained, the ultimate measure of their value must be their correlation with final performance in business. But it often requires weeks to determine a man's value on the job. If at the end of that time the tests show no correlation, much time has been lost. It is therefore desirable to use as extensive a preliminary set of tests as possible, more time being given than would be practicable in the final series. Certain preliminary checks can be made to eliminate tests that give little promise of value.

Scott suggests that persons now on the job, whose ability is known, be measured by the tests. With the general tests this is often useful. With samplings of the task the performance of one who has worked six months may be no indication of what he would have done before he learned the work.

Another preliminary check is the retesting of the same group of new people after a few days, using the same or equivalent tests. If there is no serious displacement in rank it may be assumed that the first test is an indication of ultimate performance. If the correlation between the two trials is low it is possible that a first test is of little value, unless the test itself has been vitiated by the repetition.

General Intelligence Tests Tried. When it was decided to try mental tests to aid in the selection of clerical workers in a large factory in Chicago, twenty-five new clerks were chosen by the old interview method, for the experiment. This group met three times a week for class instruction. At the beginning of each class period one or two of the following tests were given to the class:

- Courtis—Arithmetic, Series A
- Kelly—Silent Reading
- Starch—Arithmetic Scale
- Ayres—Spelling Scale
- Ayres—Handwriting Scale
- Thorndike—Handwriting Scale
- Trabue—Language Completion Tests A, C, and K
- Scott—Reading Figures
- Scott—Proverbs
- Woodworth-Wells—Easy Opposites
- Woodworth-Wells—Difficult Opposites
- Woodworth-Wells—Verb-object

Woodworth-Wells Mixed Relations
Woodworth-Wells Difficult Directions
Cancellation of Letters
Substitution of Digits for Geometric Figures
Terman-Binet Tests for Ages 10-12-14-16-18
Healey 5-block Puzzle
Münsterberg—Observation (numbers in rotation)

At the end of four months, and again at the end of six, nine, twelve, and eighteen months, an attempt was made to measure the efficiency of the graduates as actual time-clerks. Several indices were used, and the average of these taken as the true measure of ability.

It remained to determine what correlation existed between scores on the various mental tests and this index of shop success. Tests with correlations below .4 were discarded, those between .4 and .6 were considered for second trials, and those above .6 were regarded satisfactory. Most of the tests fell in the first class, a few came above .4, and almost none above .6.

Meanwhile experienced shop time-clerks were tested, but without obtaining many encouraging correlations.

On the other hand, intercorrelations were in some cases high, which indicated that the tests were reliable measures of some ability, if not of that required for time-clerks.

Job Analysis and Sampling Tried. Following the failure of the synthetic method to discover any valuable tests, an analysis of the time-clerks' task was made, to discover what factors contributed to success. A superficial study revealed the following desiderata:

1. Speed and accuracy in adding hours worked on various jobs each day in the week.
2. Speed and accuracy in computing elapsed time between time started and time finished on a job.
3. Speed and accuracy in copying data from one record to another.
4. Legibility and neatness of records.

Errors in these operations were so frequent as to require one investigator to correct the work of each ten time-clerks. It was evident, therefore, that accuracy was of first importance; and yet with two million tickets to be handled a week, speed could not be ignored.

Inspired by the Courtis drill problems, two sheets were designed, one calling for addition of time, the other for subtraction of time¹, in exactly the form required by the time tickets. These were used four minutes a day for twenty-five repetitions, and first and last scores correlated with shop success. Their value for the selection of time-clerks can be seen from results. The class was divided into quarters by scores on first tests. Shop reports six months later showed:

Best quarter—

- 2 head clerks
- 3 excellent time-clerks
- 1 good time-clerk

¹These tests are reproduced on pp. 63, 64.

Second quarter—

2 head clerks

1 excellent time-clerk

1 good time-clerk

1 fair time-clerk

Speed in Four Minutes

TEST No. 2 Name.

Errors.....

Date..

.191..

ADDITION OF LABOR HOURS

Mon.	Tue.	Wed.	Thu.	Fri.	Sat.	Total	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.	Total
8 $\frac{1}{4}$	8 $\frac{1}{2}$	11 $\frac{3}{4}$	8 $\frac{3}{4}$	8 $\frac{3}{4}$	4 $\frac{1}{4}$		8 $\frac{3}{4}$	8	8 $\frac{3}{4}$	10 $\frac{1}{4}$	8 $\frac{1}{2}$	4 $\frac{1}{4}$	
3 $\frac{1}{2}$	8 $\frac{3}{4}$	11	3 $\frac{1}{4}$	8 $\frac{3}{4}$	4 $\frac{1}{4}$		8	6 $\frac{1}{4}$	8 $\frac{3}{4}$	8 $\frac{1}{2}$	11	3 $\frac{1}{2}$	
10 $\frac{1}{4}$	1 $\frac{1}{4}$	10 $\frac{1}{4}$	5 $\frac{1}{2}$	5 $\frac{3}{4}$	3 $\frac{1}{4}$		8 $\frac{3}{4}$	7 $\frac{1}{2}$	8 $\frac{3}{4}$	5	8 $\frac{1}{2}$	3 $\frac{1}{4}$	
7	6	6 $\frac{3}{4}$	6 $\frac{3}{4}$	8 $\frac{1}{2}$	4 $\frac{1}{4}$		8 $\frac{3}{4}$	7	10 $\frac{1}{2}$	7 $\frac{1}{4}$	8 $\frac{3}{4}$	4 $\frac{1}{4}$	
8 $\frac{3}{4}$	6 $\frac{1}{4}$	5 $\frac{1}{2}$	5 $\frac{1}{4}$	8 $\frac{3}{4}$	4 $\frac{1}{4}$		6 $\frac{1}{4}$	8 $\frac{1}{2}$	8 $\frac{3}{4}$	8 $\frac{3}{4}$	6	4 $\frac{1}{4}$	
8 $\frac{3}{4}$	4 $\frac{1}{2}$	8 $\frac{3}{4}$	6 $\frac{3}{4}$	10 $\frac{1}{4}$	4 $\frac{1}{4}$		8 $\frac{1}{4}$	8 $\frac{1}{2}$	3 $\frac{3}{4}$	5 $\frac{1}{2}$	4	3 $\frac{3}{4}$	
8 $\frac{3}{4}$	8 $\frac{1}{2}$	7 $\frac{1}{4}$	8 $\frac{1}{4}$	6	4 $\frac{1}{4}$		8 $\frac{1}{2}$	6 $\frac{3}{4}$	6	6 $\frac{1}{2}$	6 $\frac{3}{4}$	4	
7	8 $\frac{1}{4}$	8 $\frac{3}{4}$	8 $\frac{1}{2}$	4 $\frac{1}{2}$	4 $\frac{1}{2}$		8 $\frac{3}{4}$	4	5 $\frac{1}{4}$	7	10 $\frac{1}{4}$	4 $\frac{1}{4}$	
7 $\frac{1}{2}$	3 $\frac{1}{2}$	8 $\frac{3}{4}$	8 $\frac{1}{2}$	10 $\frac{3}{4}$	4 $\frac{1}{4}$		8 $\frac{3}{4}$	5 $\frac{1}{2}$	8 $\frac{3}{4}$	6 $\frac{3}{4}$	11	3 $\frac{1}{2}$	
6 $\frac{1}{4}$	10 $\frac{3}{4}$	8 $\frac{3}{4}$	10 $\frac{1}{4}$	4 $\frac{1}{2}$	3 $\frac{1}{2}$		8 $\frac{1}{2}$	8 $\frac{1}{4}$	7	8 $\frac{3}{4}$	11 $\frac{3}{4}$	4 $\frac{1}{4}$	
8 $\frac{3}{4}$	6 $\frac{1}{4}$	8 $\frac{3}{4}$	7 $\frac{1}{4}$	6	3 $\frac{3}{4}$		8 $\frac{1}{4}$	8 $\frac{3}{4}$	10 $\frac{1}{4}$	6 $\frac{3}{4}$	8 $\frac{3}{4}$	4 $\frac{1}{4}$	
8	7 $\frac{1}{2}$	10 $\frac{1}{2}$	8 $\frac{3}{4}$	4	4 $\frac{1}{4}$		8 $\frac{3}{4}$	4 $\frac{1}{2}$	7 $\frac{1}{4}$	8 $\frac{1}{2}$	10 $\frac{3}{4}$	3 $\frac{1}{2}$	
8 $\frac{3}{4}$	7	8 $\frac{3}{4}$	8 $\frac{1}{2}$	6 $\frac{3}{4}$	4 $\frac{1}{4}$		7 $\frac{1}{2}$	9 $\frac{1}{4}$	6 $\frac{3}{4}$	5 $\frac{1}{4}$	10 $\frac{1}{4}$	4 $\frac{1}{4}$	
8 $\frac{3}{4}$	8 $\frac{1}{2}$	8 $\frac{3}{4}$	6 $\frac{1}{2}$	10 $\frac{1}{4}$	3 $\frac{1}{2}$		8 $\frac{3}{4}$	8 $\frac{1}{4}$	8 $\frac{3}{4}$	10 $\frac{1}{4}$	8 $\frac{3}{4}$	4 $\frac{1}{4}$	
6 $\frac{1}{4}$	8 $\frac{1}{2}$	9	7 $\frac{1}{2}$	11 $\frac{1}{4}$	4 $\frac{1}{4}$		10 $\frac{1}{4}$	6	8 $\frac{1}{2}$	6 $\frac{3}{4}$	8	4 $\frac{1}{4}$	
8 $\frac{1}{4}$	6 $\frac{3}{4}$	5 $\frac{1}{4}$	6 $\frac{3}{4}$	11 $\frac{3}{4}$	4 $\frac{1}{4}$		7	3 $\frac{1}{2}$	8 $\frac{3}{4}$	8 $\frac{3}{4}$	8 $\frac{3}{4}$	3 $\frac{1}{4}$	
8 $\frac{1}{2}$	4	8 $\frac{3}{4}$	8 $\frac{3}{4}$	8 $\frac{1}{2}$	3 $\frac{1}{2}$		9	6 $\frac{1}{4}$	8 $\frac{3}{4}$	8 $\frac{1}{4}$	4 $\frac{1}{2}$	4 $\frac{1}{4}$	
8 $\frac{3}{4}$	5 $\frac{1}{2}$	7	10 $\frac{1}{4}$	8 $\frac{1}{2}$	3 $\frac{1}{4}$		7 $\frac{1}{2}$	10 $\frac{3}{4}$	11 $\frac{3}{4}$	8 $\frac{1}{4}$	5 $\frac{3}{4}$	4 $\frac{1}{4}$	
8 $\frac{1}{2}$	8	8 $\frac{3}{4}$	5	8 $\frac{3}{4}$	4 $\frac{1}{4}$		8 $\frac{3}{4}$	4 $\frac{1}{2}$	7 $\frac{1}{4}$	8 $\frac{1}{2}$	4 $\frac{1}{2}$	4 $\frac{1}{4}$	
8 $\frac{3}{4}$	8 $\frac{1}{4}$	8 $\frac{3}{4}$	8 $\frac{1}{2}$	8 $\frac{1}{2}$	4 $\frac{1}{4}$		6 $\frac{1}{4}$	8 $\frac{1}{2}$	11	5 $\frac{1}{2}$	8 $\frac{1}{2}$	4 $\frac{1}{4}$	

Speed. in Four Minute TEST No. 3 Name. 191.

Errors. Date.

ELAPSED TIME

Started	Finished	Hours	Started	Finished	Hours	Started	Finished	Hours
7:30	11:15		7:30	10:00		7:30	11:00	
11:15	2:45		10:00	5:00		11:00	2:00	
2:45	5:00		7:30	9:45		2:00	4:15	
7:30	10:15		9:45	11:00		4:15	5:00	
10:15	5:00		11:00	4:00		7:30	4:15	
7:30	7:45		7:30	5:00		9:15	12:00	
3:45	5:00		7:45	10:30		12:45	3:00	
7:30	11:00		10:30	3:15		8:00	5:00	
11:00	3:15		3:15	5:00		7:30	4:15	
3:15	5:00		7:30	3:00		4:15	5:00	
7:30	10:45		7:30	11:00		7:30	5:00	
10:45	2:00		11:00	2:15		7:30	10:15	
2:00	3:45		2:15	5:00		10:15	5:00	
3:45	5:00		7:30	12:00		7:30	9:45	
7:30	5:00		12:45	3:30		9:45	5:00	
7:30	10:30		3:30	5:00		7:30	9:15	
10:30	3:15		7:30	2:45		9:15	3:30	
3:15	5:00		2:45	5:00		3:30	5:00	
7:30	8:45		7:30	3:15		7:30	8:15	
8:45	5:00		3:15	5:00		8:15	3:00	

Third quarter—

2 very slow time-clerks

4 dismissed because slow and inaccurate

Fourth quarter—

All had been dismissed except one, who was transferred to trucking.

Each test was scored for speed and accuracy separately. No combined score for both has proved satisfactory. The correlation between speed and accuracy can be seen from the following data:

Accuracy of Groups Selected by Scores in Speed

Best quarter.....	Average accuracy .985
2d quarter.....	Average accuracy .86
3d quarter.....	Average accuracy .71
4th quarter.....	Average accuracy .66

The extent to which each group improves with practice is shown by the average performance of each group on first and twenty-fourth repetitions of the test.

Improvement with Practice

	1st trial	24th trial
Best	23	42
2d	13	25
3d	6	15
4th	4	8

Each group after four months' drill barely attained the score possessed at the beginning by the group above. The best individual score was 50. While there was a slight displacement in rank between the first and last trials, two-thirds of the men finished in the same group they started in, and the rest moved up or down only one step, which would seem to indicate that the first trial was a fair measure of ultimate performance.

The failure of the Starch Arithmetic Test to select time-clerks points to some special ability required in computing time which could not be predicted from success in ordinary arithmetic problems, just as Courtis long ago demonstrated that ability in addition could not be taken as a measure of excellence in subtraction.

Nervousness of Applicants. Every discussion of tests for the selection of workers brings up the questions, Will not the fear of losing the job cause a nervousness which will vitiate the results of the test? And will not the average applicant be unwilling to submit to the tests?

Of course much depends upon the personality of the interviewer, and upon the way the tests are given. The use of the four arithmetic fundamentals as introductory tests served to set the applicant at his ease. The problems are so simple that he soon loses any nervous self-consciousness, but plows along with the feeling that he is master of the situation. It is often really amusing to see the self-satisfaction of an applicant whose performance is less than half the standard, but who feels he has done well because he is so sure his answers are *right!* The next two tests (addition and subtraction of time) require considerable effort, but they are followed by the Opposites and Copy Tests, which put him at ease once more so that he attacks the more difficult tests which follow with confidence. It is largely because of their value in giving the

subject a favorable start that the Courtis A Series has been retained in preference to later and perhaps more reliable measures of arithmetic ability.

Nervousness that was only temporary would disappear after several days of repeated testing. It would have been readily detected by the sharp ascent of the learning curve. Yet out of over a hundred persons observed during several weeks of testing, not more than five or six cases of stage-fright were noticed. Nervousness that was characteristic of the person at all times would interfere with business relations as well as with tests. Therefore, it is believed nervousness cannot be urged as a serious objection to selection by tests.

The applicant's attitude toward the tests is determined in large part by the interviewer. When it is explained that several positions are open, requiring special abilities of various sorts, and that it is the purpose of the tests to discover in which line the applicant is best fitted to succeed, he accepts them in the spirit in which they are really intended, not as *fault* finding, but searching for particular *abilities* which will make a valuable employee, and which cannot be discovered in any other way. Presented in this manner it is not often that an applicant shows any resentment.

A Modified Opposites Test. A time-clerk must not only be accurate, but he must also be rapid, and his speed in the arithmetic processes was only one factor.

A survey of the shop discovered some clerks writing 50 per cent more tickets than others, working under identical conditions. An analysis revealed two factors governing the volume of output: speed of handwriting and speed of making simple decisions. To measure speed in handwriting was not difficult. The next task was to find a measure of the hesitation factor.

The first test tried was the simple opposites. Given in the standard way it was open to two objections. Being oral, it required individual testing in a private room, whereas it was the aim to develop tests which could be given to a whole group at one time in the classroom. Moreover, the shop clerk is making written rather than oral responses in his regular work, so the test should likewise require written responses. The standard procedure was therefore altered to require written answers. This, of course, introduced another variable—speed of handwriting. To remove this from the equation the subject was given a second test in which he was required to copy the same words which he had just written as opposites. The significant score then was the *difference in time* between that required to write opposites and that to copy the same words.

What is really measured by this test is probably the ability to look ahead to the next operation while completing the preceding one. With more intelligent subjects there is no hesitation between words. With less intelligent ones the task of writing seems to occupy the entire attention so that the mind is not free to think of the next opposite until the last *t* of the preceding word has been crossed and the *i* dotted. And there seems to be no correlation ($-.10$, Spearman) between this necessity for attention to the task of writing and the speed with which the writing

is done. It is equally required by those who write slowly and by those who write rapidly.

A rather surprising correlation was discovered by one of the examiners, who noticed that the higher the school grade of the subject the better his score. The median retardation factor for eighth-grade graduates was 45 seconds, for those who had gone two years to high school, 32 seconds, and for high-school graduates, 24 seconds.

Correlation between scores on this test and success in the shop may be seen in the average wage increase granted to members of the four groups during a year's service.

Group	Retardation	Average Increase
Highest quarter.....	less than 20 seconds	\$3.42
Second quarter.....	20-30 seconds	2.97
Third quarter.....	30-50 seconds	2.24
Poorest quarter.....	over 50 seconds	1.61

In other words, the applicant starting at, say, \$15 a week, who tested above the upper quartile would be receiving \$18.42 at the end of a year, while his neighbor starting at the same rate but testing below the lower quartile would be receiving only \$16.61 at a year later, if he was so fortunate as to hold his job that long, which was improbable. Of those who started in the poorest group only 17 per cent survived a year, while from the upper quarter 70 per cent were still employed at the expiration of twelve months.

These two tests (written Opposites and Copying) requiring from a minute and a half to three minutes each, were added to the set.

The Difficult Directions Test. Besides speed and accuracy in arithmetic and in copying, another factor determined the value of the time-clerk, viz. his ability to make accurate quick decisions and to detect the *exceptions* in a mass of routine tickets. His instructions were complicated by countless "ifs". "If the employee works overtime, the overtime allowance must be written on a day-work ticket. But if he is employed regularly on special hours, no allowance ticket is required." "If he works on a piece-rate, use a ticket printed in red, but if he is repairing defectives use a day-work ticket printed in black on red paper." And so on indefinitely.

One clerk hesitates so long in making simple decisions that he is too slow. Another does not "stop to think", uses the wrong form, and causes the ticket to be mis-filed. Between these two extremes is the desirable man. To discover him the "Difficult Directions" test² was tried.

At first low correlations were obtained, but certain facts led to the belief that it was the method of scoring rather than the test itself which was at fault. Accordingly, a thousand tests were scored in three ways and correlated with shop success. One method of scoring gave an index of .81, and from that time on the test has been a valuable part of the set.

²This Difficult Directions Test is reproduced on page 68 with the title "Follow Instructions", which was deemed more appropriate for the type of persons being tested.

TEST No. 11

Western Electric Company
Incorporated

Name _____

Clerical Training Department

Date _____

191 _____ Time _____

Errors _____

FOLLOW INSTRUCTIONS

- 3 With your pencil make a dot over any one of these numbers, 2 4 6 8 9.
- 3 And a comma after the longest of these three words: BOY MOTHER GIRL
- 2 If Christmas comes in March, make a cross right here, _____
- 6 But if not, pass along to the next question and tell where the sun rises—, —, —, —
- 3 If you believe that Edison discovered America, cross out what you just wrote,
But if it was someone else, put the correct number in the next sentence:
2 "A horse has———feet"
- 5 Write "YES" no matter whether China is in Africa or not, _____
And then give a wrong answer to this question:
10 "How many days are there in the week?"———
- 1 Write any letter except "G" just after this comma,———
- 12 And then write "NO" if 2 times 5 are 10———
- 1 Now, if Tuesday comes after Monday, make two crosses here———
- 5 But if not, make a circle here———or else a square here———
- 5 Be sure to make three crosses between these two names of boys:
GEORGE———HENRY
Notice these two numbers: 3, 5. If iron is heavier than water,
6 Write the larger number,———
- 8 But if iron is lighter, write the smaller number———
- 11 Now show by a cross when the nights are longer: In summer?———
In winter?———
- 2 Give the correct answer to this question: "Does water run uphill?"
———
- 3 And repeat your answer here———
- 1 Do nothing here ($5 \div 7 =$ _____) unless you skipped the preceding question, but write the first letter of your first name and the last letter of your last name at the end of this line.

A year ago there was published a discussion of the Difficult Directions Test (Pintner and Toops in the *Journal of Educational Psychology*), scored according to Wooley's method. The conclusion was that the test was too easy for adults. The difficulty lay rather in the method of scoring than in the test itself. Judged by accuracy alone it is true

that the test does not give a distribution of the best fifth of all persons tested, who make scores of 100 per cent. But when the *time* factor is taken into consideration this disadvantage disappears. A still further objection to the Wooley method is that proper allowance is not made for the varying difficulty of the questions. An analysis of a thousand papers showed 3 per cent failing on the easiest question, and twenty-eight per cent failing on the most difficult one. An accuracy score was adopted which was based upon the number out of the thousand who failed on each question. The time score was merely a separation into the four groups by the median and the quartiles.

The correlation between speed and accuracy was studied in the following manner. Using the quartiles as the dividing points, all papers were separated into nine groups:

A—Rapid accurate	9	per cent
B—Rapid average	16	
C—Rapid inaccurate	0	
		25 per cent
D—Average accurate	13	
E—Average average	25	
F—Average inaccurate	13	
		51
G—Slow accurate	3	
H—Slow average	9	
I—Slow inaccurate	12	
		24
		100 per cent

If no correlating influence were at work, Groups A, C, G, and I should each include 6 per cent; B, D, F, and H $12\frac{1}{2}$ per cent; and E 25 per cent. The fact that A and I are high, while C and G are low indicates the existence of a force which tends to make rapid people accurate, and slow people inaccurate. Another interesting fact is the excess in the rapid-average group at the expense of the slow-average group, but no corresponding drift toward excellence between Groups D and F. This can be seen best in the diagram below:

		Accuracy	
Speed	A	B	C
	9 per cent	16	0
	D	E	F
	13	25	13
	G	H	I
	3	9	12

Expressed in indices of correlation (Product-moment formula) for the various groups:

A.....	.972
B.....	.009
C.....	—
D.....	.087
E.....	.214
F.....	.774
G.....	.
H.....	.129
I.....	.997

Total... .325

Used as a basis for selection, then, to obtain ten clerks from group A would require the examination of 111 applicants. Sacrificing speed but still maintaining accuracy, 10 could be obtained from Groups A and D out of 45 applicants. Allowing half an hour to each applicant this would cost 2½ hours for each clerk hired, which is excessive. Adding the applicants from Group B would reduce the total tested to 27 for 10 clerks, or less than 1½ hours for each clerk, and still obtain clerks decidedly above the average. If it is desired merely to eliminate the exceptionally poor groups, E could still be added, requiring 16 applicants to 10 clerks hired, at a cost of 48 minutes per clerk.

The value of such a selection of clerks is shown by the following table:

Group a..Best ¼ 10 clerks write 291 papers an hour, 97.5 per cent accurate.

Group b..Best 1/10 10 clerks write 246 papers an hour, 90 per cent accurate.

Group c..Best ¾ 10 clerks write 214 papers an hour, 83.5 per cent accurate.

Group d..Unselected 10 clerks write 195 papers an hour, 80 per cent accurate.

1000 papers would cost 34.4 hours by Group a

1000 papers would cost 40.7 hours by Group b

1000 papers would cost 46.8 hours by Group c

1000 papers would cost 51.4 hours by Group d

if all papers were correct. Assuming that one error costs as much as ten papers (a very moderate estimate), then the cost per 1,000 papers would become

Group a..... 35.3 hours

Group b..... 44.8 hours

Group c..... 54.5 hours

Group d..... 61.7 hours

which is a saving per 1,000 papers of

Group a..... 26.4 hours

Group b..... 16.9 hours

Group c..... 7.2 hours

Such a saving was actually experienced when selected shop clerks were substituted for untested clerks. In several departments two clerks were able to do the work formerly done by three, with a decided increase in accuracy.

Tests Must Indicate General as well as Special Ability. The modern employment manager is not satisfied when he has merely introduced the applicant to his first job in the organization. He feels that he must follow up this first placement by such transfers and promotions as are necessary to bring out all the best that the man is capable of attaining. To do less than this is not only unfair to the individual, it is also poor business. For nothing makes for a loyal force of enthusiastic workers so much as a vigorous educational and promotion program. The fact that those in high places have risen from the ranks is the strongest argument that there is a future for men of ability in the organization, and men of ability cannot be held long at any wage where they feel they are working in a blind alley.

Therefore, in a set of tests for shop clerks it is essential to introduce measures of general intelligence which will pick out the men in whom it is worth while to invest broad training with the hope that they can ultimately become understudies for executive positions. For while no correlation was discovered between general intelligence and success as routine time-clerks, considerable correlation appeared between general intelligence and success in positions of greater responsibility. Probably even a moron, if he possessed peculiar talent for that special work, might write time-tickets successfully. But the ability to organize work and to train subordinates requires intelligence which is readily recognizable thru tests.

Undoubtedly other tests might have been given, with equally good results. But it seemed wise to use standardized tests which conformed as nearly as possible to the task for which they were to select.

In a big organization most of the work of a minor executive consists in handling internal correspondence. Every fifteen minutes the mail boy drops another pile of memoranda into the "IN" basket. The department head must be able to comprehend these not-always-too-carefully-outlined requests and instructions, and make replies that the recipient can understand. Accordingly the Kansas Silent Reading Test and Trabue's Language Completion Tests were added. The Kansas Silent Reading Test was designed for school children, and its language in some cases might offend the dignity of the young men who were being tested, so slight changes were made in the wording to give it a business flavor. This modified test gave such high correlations with success on executive work (.77) that it has not been altered to conform with later revisions of the test.

Trabue's Test has two advantages. It not only gives a valuable index of general ability, correlating highly (.68) with success in responsible positions, but it also detects two undesirable types—one the man who jumps at conclusions and does not go back to check his work, the other the man who gives up too easily before obstacles. It is not safe to carry such inferences too far. Obviously it would be difficult to evalu-

ate such a correlation numerically, but the number of persons sent back as unsatisfactory in these traits who showed the same characteristics in Trabue's Test is so large as to preclude the mere operation of chance. When confirmed by evidences of the same trait in the Difficult Directions and the Silent Reading Tests it may safely be accepted as a warning of a fault to be guarded against.

Form for Recording Test Scores. Where a number of applicants must be tested each day, it is essential that the scores be presented in such a form that they reveal at a glance the standing of the applicant. For this purpose a score-card was designed similar to that used by Kitson in his measurement of students at Chicago University. A vertical column is assigned to each test and all are divided horizontally by the 25 per cent and the 75 per cent lines. The actual score is written in its corresponding column at the proper height, which makes future statistical summaries easier than the lines used by Kitson. Such a score-card classifies the applicant as either superior, ordinary, or inferior. Finer subdivisions detract rather than add to the usefulness of the card. Executives soon learn to read the records, and often send for them when considering transfers or promotions. In some cases it was desirable to make a duplicate set for the executive's file.

Individual Cases. How the definite information furnished by tests can supplement the general impressions of the interview, the following individual cases may illustrate.

"M" was a graduate of the University of Chicago. That she possessed a keen mind no one who talked with her five minutes could doubt. An English college woman had just achieved wonders in the cost accounting department, and was calling for an assistant. "M" was given the job. Her first task was to learn to compute percentages on a slide-rule. After several days of patient effort her superior appealed for an exchange. "M" could not do the work. Still she was recommended as unusually intelligent, willing, with pleasing personality. She was sent to be tested. The results were extremely striking. Below the lower quartile in all arithmetic tests, she was almost without equal in the excellence of all other scores. A vacancy was created for her in a correspondence group, and in a short time she was head of the group, auditing all correspondence which went out and dictating many letters herself from technical data furnished by the engineers. By the use of tests in the beginning she might have been saved the discouragement of the first failure, which might under less favorable conditions have lost a valuable employee to the company.

"W" was a high school graduate, nineteen years old, who was promoted to a position where he was required to check the typewritten purchase-orders for tools against the engineer's longhand originals. He allowed an error to pass which cost the company several thousand dollars and kept the U.S. government waiting ten days longer for its equipment. Under the old régime he would have been discharged at once. Instead he was sent for test and placement. The tests showed him to be rapid but inaccurate. He was placed where speed was the main requirement, and where his errors could do no harm, and has proved highly satisfactory.

The Demoralizing Effect of Misplacement. The importance of proper placement as a saving to the employer is obvious. But the value to the individual worker must not be overlooked.

The demoralizing effect of failure can hardly be overstated. The man who gets into the wrong job deteriorates so rapidly that unless he is a strong man he is liable to be permanently dwarfed. Like the bad-fitting shoe, the bad-fitting job causes all manner of painful deformities. One has only to sit behind the application desk of an employment department for a few days to see its havoc written into the souls of men. Discouraged, disappointed, restless, stunted, they drift from job to job, their precious youth wasted as mere servants of the weekly pay-envelope, their children starting the race of life with the handicap of poverty upon them, they finally find a haven as janitor or elevator operator, from whence they who survive finally become society's pensioners. A dark picture, to which there is a brighter side, for if "weeds are but flowers for which man has not yet found a use" then failures are but workmen who have not yet found the right job. To find the right jobs for these failures is a task which has its own rewards.

"R" was such a failure. He was eighteen years old, and had been newsboy, telegraph messenger, delivery boy on a department store truck, "rustler" in an express office, and now finally stock-clerk in a big factory. With no definite aim, never satisfied nor satisfying, holding each job a few months, then hunting another, anything so long as it paid enough money, he was on the high road to becoming a chronic failure. Nothing had ever really aroused his interest. Fortunately, instead of being "fired" he was sent to be tested. The scores were interesting. Ability was there without doubt. But in what useful line? Terman-Binet tests showed a good normal adult, but in arithmetic he was low. (He left school because he could not master decimal fractions.) He thought he would like to be an electrician, so a job was found for him in a motor repair shop. Here he was fortunate in finding a sympathetic teacher who encouraged the boy to do his best. Soon he was given outside jobs. He showed a liking for climbing, so crane and elevator jobs were passed to him. Now he is an experienced motor maintenance man, with real money in his pay-envelope, and two years of steady service safely behind him. He has found his job.

"G" was another failure. He had been tried on various tasks, without success. Tests showed more than normal ability. Nineteen years old, he was still drawing the pay of an errand-boy. He had been a beginner for five years, and had "gone stale". It was decided to arouse his ambition by giving him a big promotion. He was made assistant to the clerk who ordered raw material from the blue-prints, at nearly twice his former wage. He entered evening classes in drafting and shop practice, and applied himself to his work with a will. In six months his whole personality was altered. He spoke with assurance, held his head up, walked as though he were going somewhere, and had in every way "made good".

The sense of accomplishment is one of the strongest incentives to drive a man to do his best work. It will arouse effort which no wage

system ever can call forth. The schools are beginning to recognize the importance of this motive, and have in some instances remodeled their machinery so that the slower moving pupils can take three years to reach the goal their more rapid fellows reach in two. The slower pupils are thus saved the discouragement of a "not passed", and are trained for success rather than for failure; while the rapid pupils are spurred to effort comparable with their ability, and so saved from mediocrity.

Conclusions. 1. Highly specialized tasks may require particular ability which is discoverable only by tests designed for that purpose. For such tasks general intelligence tests are of little value. When tests for the special ability are perfected, the persons who are selected by these tests are much more efficient than unselected workers on the task, making it possible to accomplish more and better work with a smaller force.

2. General intelligence tests are increasingly useful, the higher in the organization scale they are applied. Their value is threefold: first, as a check upon special tests in the selection of applicants for primary tasks; second, as an indication of ability to ascend the ladder to more responsible positions, when promotions are to be made; third, to indicate whether failure on a particular task was due to lack of special ability for this work, but with promise of success in some other line, or whether it was caused by giving too much (or too little) responsibility.

3. Certain factors of success, which make up what is commonly called "personality", are not measurable by test, but must be judged by an interview. Selection cannot therefore be delegated to a routine clerk, but must always be made by a trained observer. There is danger that too sweeping inferences may be made from test data, which will be unfair to the applicant and to the employer.

4. The observer must be on the alert to detect evidences of ability or disability not discoverable by routine scoring of the tests.

5. The employment manager can be greatly aided in making placements if he can have access to the applicant's school record in terms of standard scales. To know the spelling and grammar ability of an applicant stenographer, to know the arithmetic ability of an applicant book-keeper, would save much waste to the applicant and the employer.

6. By coöperating with the university psychologists, employers may save much time in designing tests and avoid the danger which follows when over-sanguine amateurs attempt to use tests whose significance they do not understand. Many so-called tests are being used at the present time in a way that can result only in injustice to the applicant and loss of valuable men to the employer. The Bureau of Salesmanship Research at Carnegie Institute of Technology is a notable example of what can be done when employer and university get together.

7. The employer who has once been so fortunate as to obtain and use a really scientific set of tests will never again be satisfied to pick his men by the haphazard impressions of an interview alone.

Intelligence Test of High School Seniors

OSCAR H. WILLIAMS, State High School Inspector

THE state-wide intelligence test just now commanding the attention of senior classes in Indiana high schools is perhaps the first attempt in the history of mental measurements to apply a test of general intelligence to a group extending over an entire state. The test employed is known as the Indiana University Intelligence Scale No. 1. The scale is a series of tests of general intelligence devised by Dr. and Mrs. Pressey under the guidance of Professor Book. They are derivatives of the Simon-Binet Tests of intelligence and are consequently based on the same general conception of the nature of intelligence as those well-known tests, viz. *that human intelligence consists of the sum total of the thought processes which constitute mental adaptation*. They are tests of native ability to combine ideas and adapt the combination to a given purpose.

The Indiana Scale was developed in connection with certain problems of research carried on thru the departments of psychology and sociology of Indiana University in the autumn of 1917. The problems were combined into a single research having a twofold aspect: first, that of selecting gifted children in the public schools, for special classes, extra promotion, or other assignment; and, second, that of determining the proportion of mental defectives among the school children of a certain county. The scale was later used in testing the pupils in three representative high schools of the state, besides several hundred high school children in other parts of the country. The utility and applicability of the scale have thus been firmly established.

The present test is an initial attempt to apply the Indiana Scale under approximately uniform conditions to a fairly homogeneous group,—the senior students in the six hundred commissioned high schools of the state. The test is administered under the direction of the division of inspection of the State Department of Public Instruction, in coöperation with the department of psychology of Indiana University. Its primary purpose is to discover pupils of superior mental endowments, thereby making it possible to give these students special help and encouragement to continue their education. Possible uses as a basis for vocational direction and placement of the youth of high school ages are yet to be determined. The experience of the Army Personnel Office in discovering and utilizing talent in the training camps during eighteen or twenty months of service by means of similar tests suggests fruitful potentialities in this direction.

Early in the present month, a letter announcing the proposed test and urging coöperation in its administration was sent to the high

school principals. A small folder, explaining the value and significance of mental tests and suggesting possible uses of the present test accompanied the letter. Interest and enthusiasm on the part of the principals in the coming test were spontaneous and supreme. Many wrote urging that materials be sent without delay; others requested the privilege of giving the test, not only to seniors, but to the entire school.

Test materials have this week been sent directly to the schools. These include individual test papers for every member of the senior classes, an Examiner's Guide containing full and explicit instructions for giving and scoring the tests, and a report form for compiling individual scores and other data for every school. The principal and a teacher or teachers designated by him will have personal charge of the test in each of the schools. Not the least of the benefits that may accrue to the schools will be the initiation of hundreds of high school teachers in the use and scoring of mental tests.

All test papers and a summarized report of individual and class scores will be returned by each school to the State Department. From the returns a combined report showing the distribution of intelligence, the intelligence rating of the various schools in widely separated communities, will be compiled. An intelligence map of the state is among the possibilities.

The tests are ten in number, constituting a series testing a wide range of mental capacities from simple rote memory to practical information and use of analogies. Each test in turn involves some twenty mental operations, carefully graded, passing by easy stages from simple to more and more difficult thought processes. Probably not more than one pupil in every one hundred will possess the mental ability to make a perfect score. A time limit ranging from one-and-a-half to four minutes is placed on each division of the test. The entire series will require about forty minutes to administer.

The tests are so designed as to require a maximum of thinking and a minimum of writing or clerical effort, the answers consisting of a single word, a line beneath or a cross beside a printed word, or at most the re-arranging of words in logical order in a sentence. Thus they satisfy the requirement of economy of time and effort both in taking the tests and scoring results.

As suggested above, the essential purpose of the test will be to uncover and mark native intelligence or mentality, which may or may not be greatly influenced by school training. A secondary object will be to discover what correlation, if any, exists between such native endowments and various factors in the pupil's environment. Provision is made for obtaining in every case data bearing on such factors as the pupil's social and economic status, as suggested by the father's occupation and yearly income; the pupil's scholarly proclivities,—his favorite subject in high school, his scholarship rating in the previous (or junior) year, his college aspiration, his vocational or professional outlook; and time spent and course pursued in the high school.

Such in brief outline is the plan and the scope of the projected test. Of what significance to the educator is such a testing of high school

seniors likely to assume? What educational implications or results may it be expected to yield? The answer has already been suggested in part. Its primary object is to locate superior students to the end that such students may be urged and helped to go on to college. One such student given proper training may prove of priceless value to society. One of the tragic wastes of our society is the frequent neglect and loss of genius from sheer want of educational opportunity.

The intelligence rating of every pupil taking the test will become a constituent part of that pupil's high school record. His rating or intelligence score will thus accompany the pupil's certificate to college or employer. Of this important fact all are informed at the time of taking the test. In this way, college faculties will be furnished in advance with a clue to the native ability of their student material. They can determine without delay whether or not a given student is doing work in college in keeping with his native capacity.

Again, the proposed testing is likely to give to the schools some insight into the relative abilities of their pupil material. It may afford some clue as to the degree in which a given school is securing results in relation to the abilities of its students.

Other and subsequent uses are suggested by the possibilities in giving vocational guidance or direction, assistance in dividing classes into sections according to ability, and developing types of courses of instruction suited to varying degrees of mentality.

Scientific Curriculum Construction

W. W. CHARTERS

The Factors in Curriculum Construction. The curriculum is a fusion of three ingredients: life objectives, human achievements, and child nature. Each of these possess certain characteristics which are essential to each and which influence the product.

Life objectives are those purposes which direct and control the lives of men. They constitute the aims of action for which men strive and have been commonly designated in pedagogical literature as the aims of education.

A determination of the aims of education is a necessary factor of curriculum construction because without them selection cannot be made from the great mass of human achievements which are available for instruction.

Human achievements represent the advance in control and ideals which man has made over instinctive control and unconscious purposes. These controls and ideals have in a very fundamental sense been achieved by man, bought with the price of thought and patience and effort. And by their transmission to ensuing generations, conservation and progress are alone possible; without this transmission the young would immediately sink back into lower than primitive conditions and the laborious efforts of our ancestors would disappear in a single generation.

These human achievements are multifarious. They consist of inventions in mechanical appliances, creeds in religion, customs of social life, principles of science, technique of art, habits, ideals, and prejudices—all the inheritance from older generations. Everything that is given to one generation by its forbears, or added by itself, except its instinctive controls and physically based cravings, have been achieved by man and are potential elements of a curriculum.

In the processes of time these achievements have been organized, in part. In part they await organization and in part defy it. The organized subjects are known as the sciences, literature, history, art, political science, economics, sociology, and so on. The unorganized achievements constitute all the remainder of experience, the margin out into which the boundaries of the organized area are being constantly pushed as new subjects take on form and system. And outside of the limits of possible organization lie a great mass of miscellaneous achievements, disconnected, accidental, and local.

From this mass of achievements selections for the curriculum are made upon the basis of the aims of education, on the one hand, and the psychological characteristics—interests, needs, and abilities—of children, on the other. The child must live his immediate life and at the end of

his youth enter upon the participating and contributing adventures of adult society. These abilities, interests, and needs determine the order in which he is made acquainted with human achievements. They, also, determine in part the selection of the achievements; for we cannot assume that the achievements which are most useful for adults are, also, most necessary for children. Adult living does not value highly for its purposes such achievements as infant games or fairy stories altho it is clearly evident that for children they perform a notable service.

The foregoing statements are common property. The relation of the aim of education to the selection of subject-matter has been recognized in the writings of the great educators since the time of Plato, and their relation to the nature of children has been recognized with growing clarity since the time of Rousseau. They have all been discussed for generations and are today the object of widespread discussion.

The Causes of an Inadequate Curriculum. But we are not satisfied that the curriculums of the schools fit the needs of modern life, and our discussions do not seem to produce adequate improvement.

That this is the case is due to three factors which are worthy of careful consideration. The first is the assumption made universally until in recent years and by many at the present time that the best method of equipping for life is to acquaint the young with the organized fields of knowledge in such order as to reveal the systematic arrangement of each. The second is the failure to analyze the aims of education—continue the analysis until objectives sufficiently detailed have been determined. The third is the failure of educational psychology to assist the trial and error method of testing, by providing a body of facts according to which the material of the curriculum may be arranged in proper pedagogical sequence.

The Organized Fields of Knowledge. There is much to warrant the assumption that the organized position of human achievement such as the sciences, literature, and history constitute the material most useful for the fulfillment of the aims of education. It is the product of master minds and of painstaking followers who with admirable singleness of purpose have devoted their lives to the discovery and systematization of their fields. The influence of their substantial findings upon human life is often spectacular and seldom negligible. Then, too, the material seems easy to learn because of its orderliness and easy to impart because of the system. Further than this, the assumption has gained and held credence that when principles are mastered the learner himself can make their applications to his purposes and problems of living. These beliefs give plausibility to the assumption that the best that man has accomplished is the most nearly adequate equipment for the rising generations.

The Best That Man Has Accomplished. But the validity of the assumption rests largely upon the interpretation of the meaning of the "best that man has accomplished". For it may mean either the most highly organized, the most original, the apical achievements, those which by their sheer sublimity make man as the master mind of the world, or

it may mean the most useful where the amount of utility is determined by the number of people who can profitably use it and the value it possesses for them. It cannot be assumed that the best examples of mental brilliancy are the best instruments for daily living. As an example of the brilliance of the human mind, the discovery of the law of gravitation immeasurably outshines the discovery of cooking; but on the basis of daily living, the instruction in cooking is vastly more important than the teaching of the law of gravitation. If, therefore, by the term, the best that man has accomplished, is meant the most useful for daily living, the assumption that the organized fields of knowledge are the most nearly adequate equipment for the immature does not hold.

That acquaintance with these organized masses of achievements is not necessarily the best material for instruction is due to the method of their growth. It is a matter of common knowledge that the older sciences and probably all the newer ones as well had their origin in the context of daily life. Many were begun by the common people. The women of the tribe searched for herbs to cure tribal ailments and thereby introduced the study of botany. And in due course of time, when the tribe could afford the luxury of supporting a doctor, the medical profession greatly widened the search for herbs and collected information about and samples of new and potent plants discovered in the far parts of the earth. As time went on it was only natural that the inquisitive mind of man should seek to classify all the plants of the world, whether useful for healing or not. Later when the microscope and the theory of evolution were discovered new generations as they studied plants became interested in the relation of one plant to another, in the structure of plant cells, in the relation of plant life to physics and chemistry, in microorganisms, in adaptation to environment, and the other multifarious problems of modern botany. And as they discovered new facts, the botanist sought for basic principles and organized his knowledge in as systematic and logical an order as possible. But the study of plants in their relation to the art of healing has long since been outgrown and is now relegated to a minor division of the field which is studied by druggists and doctors.

The content of highly organized fields of knowledge is essentially not determined by daily needs and practical considerations. The specialist in the pure sciences has only secondary concern for the practical utility of his discoveries. He leaves to the less well organized applied sciences the selection from his products of those which may have practical utility.

But while he is not influenced by his neighbors in practical life, he is notoriously sensitive to advances made by his neighbors in other highly expert fields. It is a matter of common knowledge that when biology substantiated the theory of evolution and worked out a technique, it profoundly influenced ethics, religion, and sociology.

The results of this are twofold. In the first place, the content of a field of knowledge is partly accidental. It is obvious that discovered knowledge must affect the discovery of new knowledge, but there is no great guiding power which gives research constancy of direction; no

scientist can foretell what the content of any field of knowledge will be in the next generation. It is accidental because of the personality of investigators and the discoveries in neighboring fields. Darwin was a man of a certain patient and persistent type who took hold of a theory floating in the intellectual atmosphere of his time and proved with adequate rigidity that it was correct. If, however, he had been of another type, a speculator and a dreamer, he would have been content to talk about the theory, and today the content of many fields of knowledge might well have been quite different. The discovery of radium was not a necessity. Curie might have been run over by a cab twenty years earlier, and if so the content of physics and chemistry would have been substantially changed from what they now are. Perhaps at some time and under some circumstances, these discoveries would have been made but there is no certainty of such an occurrence. Every science is filled with undeveloped leads which, had they been developed by some investigator, would have made profound modifications in its structure.

In similar manner the development of a field of knowledge depends upon the discoveries in other fields. The discoveries of science have affected the content of theology and ethics, those in chemistry have influenced biology, and those in neurology have affected psychology. For the alert worker in one field narrowly watches the methods in another field, and when he discovers others which seem to be fruitful he is almost as likely as not to drop the problems or the methods he is working with and take up the new ones which seem to him to be more interesting and significant.

So for these two reasons the content of any field of knowledge is not, as the popular mind supposes, a great immutable structure based upon and inherent in the metaphysical nature of matter. Rather it is constantly changing and to a very considerable degree is dependent upon the interests and previous discoveries of quite mutable men.

We cannot foretell the social problems of the next generation with any greater accuracy than we can prophesy about the content of science at that time. For instance, if wiser counsel had prevailed in the conference of the War Lords in July, 1914, and it had been thought better to wait ten years or twenty before the campaign for a place in the sun should be launched, we should have had problems of daily living in America today notably different from those which now confront us. Whatever our ideas about predestination and purposes which shape our ends, it is agreed to by all except the extreme fatalist that the problems of daily living change from day to day and from generation to generation according to no discovered laws which are capable of controlling their emergency.

It is evident, therefore, that the content of the fields of organized knowledge will not parallel the problems and needs of daily life. The fundamental discoveries of science were not obtained in a search for solutions of the fundamental problems of living. They were obtained by men who were seeking to develop narrow fields of knowledge whose content is to a considerable degree influenced by chance. Indeed, if science had developed with the avowed purpose of solving problems of daily liv-

ing, only the most nimble gymnastics supplementing the most sensitive attention to the changes in daily thinking and action would guarantee that the subject-matter could be accepted *in toto* as an adequate equipment for living.

The Function of the Applied Sciences. Outside the school, this lack of connection between the fields of knowledge and practical life has been recognized by both the scientist and the practical man, and as a result the new, less highly organized and less sure-footed applied sciences have been developed. The purposes of these agencies is the study of both the fields of science and the field of practical life in order to make applicable to the latter, not all the facts of science, but such of them as will be of use in the solution of practical problems.

This plays ducks and drakes with the logical organization of the pure sciences. For with an audacious nonchalance the practical science takes the facts and principles as it needs them. Its problem is to satisfy some human need and in doing so it may select facts from a dozen fields of knowledge and organize them, not on the basis of the principle of each field but according to their utility for meeting practical needs.

The methods of selection in the field of applied science are significant. The applied scientist may do either one of two things and usually does both. He may take up the facts and principles of science one by one and ask the question, Is this of any practical use? If it is he develops the method of making it usable. If it is not he discards it. Or he may study the problems of practical life, and, finding some of greater importance than others, he may ask the question, Is there anything in science which will help in the solution of these problems? And in similar manner he selects from the sciences those items which he can use and ignores the others.

This procedure is significant because of the ruthlessness with which it treats the sciences and the sensitiveness it shows for practical needs. Daily living is paramount and the facts of science are merely collections of tools which have varying degrees of usefulness. This is particularly significant in curriculum construction because it is an example of a second method of constructing a curriculum. The first we have been discussing from the outset—the determination of the items in a curriculum by the assumption that the aims of education can be best met by the impartation of essential content of each of the highly organized fields of human achievement. The second, that of studying the daily lives of men to find the most important problems and the selection from any source available of such items as will solve these problems is the method of applied sciences.

The Case of Elementary Science. The case for the organized fields of achievement is still weaker when applied to the teaching of immature students. For to meet this immaturity it has been assumed that the outline or skeleton of an organized body is easier to learn than is the more fleshly clothed structure. This accounts in part for the fact that books for beginners are thin while books for the initiated are fat. But a moment's consideration will show that this procedure is wrong. In such a case volumes and chapters are condensed into single sentences and

paragraphs, the material is stripped of the content by which alone it can be made clear. One by this means obtains a faithful reproduction in miniature of the whole field, but the immature and inexperienced mind is unable to grasp it and therefore to use it. The psychology of the child teaches this most convincingly that children achieve the ability to organize only to a very slight degree. They are not specialists either with respect to scientific interest or to the patience necessary to systematize logically a mass of unrelated facts particularly if it is of little interest to them. They are like most people who are not specialists, the creature of their more immediate needs and pressing problems.

Consequently, the curriculum which is presented to children as a reduced photographic reproduction of a great field of knowledge neither solves the problems of their present or future daily lives nor does it obtain their comprehension. To a very considerable extent it is a meaningless recitation of facts lightened here and there by occasional spurts of interest in isolated items. This accounts for the commonly made observation that the sciences from which so much was expected when they were introduced into the schools may easily become as formal as any subject in the curriculum. The students of these fields of knowledge are often "like birds which fly abroad to forage for grain and bring it home in their beaks, without tasting it themselves; to give it to their young".

The Failure to Analyze the Aims of Education. The second factor which operates to make curriculums inadequate for daily living is the failure to analyze life objectives in sufficient detail. The meaning of this can be made clear by illustrating the method of analyses of life objectives and determining what detail is sufficient.

The aims of education are statements of the best objectives for which men strive, and these have been discussed for centuries. Some thirty-five statements are found in a history of education and they divide themselves into eight classes: morality, religion, knowledge of classics or environment, social leadership and efficiency, discipline, development of the powers of the individual, complete living, and practical success.

Among investigators, Plato was the first to discuss the aim of education and determine the curriculum. His aim is to develop knowledge of the good. His curriculum which he presents in some detail he attempts to justify in terms of his aim. And he asserts its content to be music—Dorian or Phrygian; mathematics—arithmetic, geometry, and astronomy; literature—tragedy (always containing only images of the "good") but no comedy; rhythm; harmony and observation. But he makes no detailed attempt to determine the sequences of his items according to the nature of children.

Twenty-two centuries later Spencer defined the aim of education as complete living. He analyzed this into five types of activity: self-preservation, earning a living, parenthood, citizenship, and occupation for leisure, in decreasing order of importance. From these he deduced a curriculum consisting of physiology, reading, writing and arithmetic, physics, chemistry, biology, sociology, ethics, and psychology. But, like Plato, he was not concerned with the sequence of the items of these subjects.

Rousseau in the centuries between Plato and Spencer defined the aim of education as follows: "The aim of education is to develop naturally the innate instincts and interests of an individual in such a way as to lead to an expansion of these natural powers in a nature which is given, not made." His curriculum deduced from this aim consisted of subject-matter derived from nature (not man), the science of duty, language, Robinson Crusoe, manual labor, and so on. In addition, he described, as Spencer and Plato do not, his conception of the psychological nature of the individual and by it arranged the sequence of the items in his curriculum.

These three samples of curriculum construction display the failure to analyze. Each mentions certain subjects which are to be taught. But what physiology and chemistry should be taught for complete living? What subject-matter derived from nature is best to develop the natural powers? What Dorian music is needed to develop ideas of the good? The answer is not given. Possibly it is assumed that these fields of knowledge should be presented in organized form altho it is evident from the foregoing discussion that such an assumption is invalid. More probably the great educators and all writers on the curriculum have not been concerned with details and have left to teachers the selection of the parts of the subjects to teach.

And herein lies the difficulty. These aims of education are stated in very general terms so vague in content that they provide no standard for selection of subject-matter. What is needed to supplement them is a method by which the aims may be brought into contact with detailed units of subject-matter.

The same difficulty appears in teachers' meetings wherein the claim is constantly made that the curriculum should fit the needs of children. But what the needs of children are have not been defined and no method for determining them has been perfected. And as a consequence the most enthusiastic advocate of adaptation to needs is more likely than not to preach reform and practice conservation.

This great gap between the aim and subject-matter, which cannot be bridged without analysis, produces some rather interesting results. It has been found that the great educators who have different aims frequently select the same subject-matter by which to realize it. Nor is this always due to logical results logically arrived at, but rather to the fact that the gap is so broad that logic cannot leap it and the writer falls back on tradition and teaches what is in the schools. Sometimes it is due to the personal opinion of the writer, and this opinion is frequently formed upon the basis of his own personal experiences as in the case of Rousseau, who bridged the gap by expanding the informal, out-of-school, curriculum thru which he as an orphaned waif was conducted.

Bridging the Gap between Aims and Subject-Matter. However, it is possible to bridge this gap by a much more scientific method and arrive in the zone of activity within which objections and subject-matter mingle and face. This method is the analysis of life objectives.

An illustration will make this clear. Spencer has defined the aim of

education as complete living and has said that certain subjects as physics and chemistry will constitute the curriculum. But we are at once faced with the question, What is complete living? Fortunately for us, the author analyzes it into self-preservation, earning a livelihood, parenthood, citizenship, and occupation for leisure. But he leaves the matter there and in leaving it thus, he helps but little. If one of these—earning a livelihood, for instance—is taken, the question must be repeated. What must you know to earn a livelihood? Obviously, an answer to this cannot be given in that form. For it depends upon what form of livelihood is selected. Analysis is again needed to determine the objectives of livelihood. This will reveal among other things a list of possible occupations, as for instance carpentry. Again the question must be asked, What does a carpenter need to know? This involves a further analysis of the “jobs” a carpenter has to perform, the skills he needs, the ideals that should dominate his work. Suppose the jobs are listed and that one of them is building a frame dwelling. Again analysis is needed to determine what are the subordinate jobs in this larger job. These consist of framing, shingling, and so forth. Selecting shingling as a job, again the question is raised as to what processes are necessary in shingling. These consist of the selection of the shingles, their treatment, method of laying and nailing. Again if we select nailing as a process, it may be necessary to analyze nailing into its processes as the most convenient nail receptacle, the finding and placing of nails, and hammering.

It may seem a long and tedious stretch to travel from complete living to nailing shingles. Complete living is a fine-sounding phrase and shingle nailing is very prosaic. But the steps are logical and the distance must be traveled by the curriculum maker. For the fact cannot be successfully controverted that while indefinite and general aims of education have had great influence upon curriculum, use has been negligible except where the enunciator of an aim or his adherents have taken the trouble to work out the details of the curriculum thru some analysis of the aim.

This analysis is tedious, and educators who have been in the habit of settling all problems of education including the curriculum by argument rather than by investigation do not hasten to undertake drudgery. In this they are, of course, quite different from the scientist who spends hours and days upon the details of some small item of investigation and who often has to discard weeks of work because some slight error has crept in. But as time goes on and the scientific attitude pervades educational investigations more widely, this drudgery will be cheerfully undertaken and then but not till then can we hope to substitute established fact for personal opinion.

The Limit of Analysis. Two factors determine the extent to which this analysis needs to be carried on. The first is the nature of subject-matter, the second is the ability of the learner.

The subject-matter of education, which is constituted by the achievements of man, is a collection of small units which may adhere in larger groups and these in still larger groups. For instance, to take the illu-

stration used just above, carpentry is a collection of projects such as floor laying, framing, shingling, etc., and these projects are groups of subordinate processes as for instance where shingling consists of units such as driving nails and laying shingles.

Working Units of Achievement. The significant fact is that the unit of achievement in this case is not carpentry and not shingling but driving nails or laying shingles. By this I mean that man in his attempt to achieve the conquest of his environment worked in very small units. He covered one short step at a time rather than the whole distance with seven-league boots. His total progress was made by the achievement of one tiny objective after another. This is true of carpentry, which is not a unit but is a collection of many units. This is equally true of the great organized fields of knowledge. Science has been developed bit by bit. It is obviously true of history which consists of a series of acts. Just as clearly it is true of the great life objectives. Comenius's "eternal happiness with God" is a collection of acts dominated, to be sure, by a single purpose but nevertheless achieved by minute actions carried thru from moment to moment.

The fact that all achievements have been made in small units is of supreme importance because it determines the form in which man uses his achievements when made, the form in which we find them recorded and, therefore, the form in which he transmits them thru instruction. This is the level on which instruction can most economically be given. The analysis of life objectives must, therefore, be carried on until it reaches the zone of man's working objectives. And the curriculum when constructed will consist of a mass of these working objectives which for purposes of easy handling may be collected into groups called subjects, upon the basis of similarity of material or principles. Perhaps the chief difference between the curriculum of one who has not attended school and one who has, lies in the fact that the former has a mass of unorganized units while the latter is able to see principles running thru certain groups of the units. But whether the curriculum is organized or not its elements are these working units. Nor can I refrain from saying that if the principles are not used as integrating instruments in classifying units otherwise existing as isolated facts in the experience of the student, they are quite useless except for decorative purposes.

This idea of working units of achievement has not received the attention it merits in curriculum making. It is sometimes recognized and used by the teachers after the curriculum is made because the teacher has to use it if the children are to learn anything, particularly in the lower grades. But even there and much more constantly in the upper grades the textbook makers and the teachers give the material to children in units which are too long, which are longer than those which man uses and teaches outside school. The illustration of neatness comes to mind. Teachers say to children "Be neat" or "That is not neat"; but the child may not know how to achieve neatness. Certainly man reaches it by smaller steps. Outside school in the home it means such jobs as picking up articles on the floor, arranging books in regular

piles, and hanging coats on appropriate hangers. Therefore, the teacher's "neatness curriculum" must come after an analysis of neatness and the listing of the objectives which when reached will produce neatness—such as no scraps on the floor or an even left-hand margin.

And in similar manner teachers, superintendents, and professors of education talk about social efficiency, complete living, morality, and the other current aims of education, but the failure to analyze these and re-analyze them until the working objectives have been reached has resulted in two worlds, one world of talk about curriculums and another world far removed from this, the teacher's daily curriculum, still controlled by tradition and habit and but little influenced by any aim that resides in the other world. The gap between the two worlds can be bridged and by a thoroughgoing analysis of the aims of education.

Working Objectives Lie in Zones. The working unit of achievement constitutes the theoretical limits of analysis. In theory, analysis proceeds until the working objectives are reached. In practice, however, it is found that these working objectives lie in zones. And the degree of analysis after the upper boundary of a zone is reached is dependent upon the ability and previous experience of the learner.

For instance, the federal government has for the last two or three years advocated the canning of tomatoes by the cold pack method. It is evident that for some women no further information is necessary. Stimulated to can tomatoes by a patriotic appeal they proceed to do so because they had the experience and knew how to do it. But some women who accepted this objective as worth-while did not know how to do it. And for these analysis was necessary. The steps in performing this consisted of collecting the canning utensils, taking care that everything was clean, testing the jars, lids, and rings, grading the tomatoes, scalding them, packing in jars, flavoring to suit taste, sealing lightly, cooking for twenty-two minutes, and sealing tightly. This evidently might be in sufficient detail for some experienced women with less knowledge than the group first mentioned. But as a matter of fact it was found that the demonstrators in their canning campaigns had to analyze each of these directions into still smaller steps in order to reach the working units of the practical housewife. For instance, instead of saying merely, Collect the canning utensils, taking care that everything is clean, some demonstrators found it necessary to give these steps: Secure a well-screened, light, airy room, free from dust and protected from drafts, containing a good stove, a supply of clean fresh water, a table, a dishpan, a kettle with a lid, teakettle, paring knife, sack of salt, teaspoon, tablespoon, supply of glass jars, etc., and wash the hands thoroly, clean finger nails, and don clean apron and cap. Moreover, it is evident that with still less experienced cooks, with little children for example, some of these steps would need to be re-analyzed as, for instance, how to wash the hands thoroly or don an apron. It is evident also that this canning project would be too large a unit for very young children who would not have a mastery of the items obtained by the final analysis.

It is, therefore, clear that the limit of analysis is the zone of working objectives within which lie the units of achievement with which man

works, that above this zone the analysis of aims of education must not stop, but that within it the degree of analysis depends upon the experience and ability of those who are receiving instruction. The failure to conduct the analysis of aims of education from their lofty and nebulous heights thru intervening levels to the zone of man's working units of achievement has contributed in no small degree to the static condition of the curriculum presented to the children by classroom teachers.

Collecting Achievements. Once, however, this has been done, once the specific objectives of daily living have been deduced from the major objectives of life, the next step in curriculum construction follows. This is the collection of the best method man has discovered and used for the reaching of these objectives. When the specific problems of daily living are determined it is then possible to select from men's achievements those practical solutions which have been found to be most satisfactory. And again it should be repeated that these are not necessarily found to be identical with the highly organized fields of knowledge. They are found in parts of these fields, to be sure, but they are also found in the less highly organized fields, in the portions of the field capable of ultimate organization, and in those portions which defy organization. They are appropriated wherever they are found without regard to any standard of origin and entirely upon the basis of their usefulness.

Educational Psychology. The third factor which has affected the progress of curriculum construction is the failure of educational psychology to render adequate assistance in classifying or delimiting the interests and abilities of children in such detail as to provide directions for sequence in instruction grade by grade. Whether this task is incapable of performance we do not know at present, but that it has not been done is clearly evident.

The problem in curriculum construction which educational psychology might be able to solve is an important one. We have seen that the analysis of the aims of education and the collection of the best methods of fulfilling them gives us a mass of material but it does not provide any basis for the determination of the time of teaching this material. If educational psychology could furnish information sufficiently definite for the curriculum constructor to determine objectively that one item or another should be taught in one grade or another, the assistance would be vital.

Stages and Development. Some attempts have been made at a classification of the years of childhood according to stages, but these are not based on scientific ground, there is no consensus among the writers themselves, the stages are usually two or more years in length, they are recognized by the writers as at best only approximations, they do not conform to grade units, and are not accepted by teachers as being practical. Some attempts, also, have been made to obviate the difficulty of analysis of the short complex life of the child by throwing it on the screen of history in magnified form thru the recapitulation theory, wherein the claim is made that long stages of evolutionary development are represented by short stages of child life. This conception has been

carried farther in the culture epochs theory which seeks not only to determine the stages but also to solve the problem of curriculum making on the basis of the claim that in each stage of child development determined by the recapitulation theory, the culture of the corresponding historical stage should be the subject-matter of instruction. But both of these theories seem at present to be untenable.

However valuable educational psychology has been for methods of instruction, its contribution to curriculum making, particularly to curriculums suitable for teaching children in groups, has been almost negligible.

Education and Life. That there is a necessity for a more thorough testing and tabulating of the interests and abilities of children is evident from the current beliefs about the relation of needs of children to the curriculum. There are, in particular, those who hold that the curriculum should be determined by adult objectives to the extent that children should be taught the solutions to problems which they will meet in adult life. There are others who hold that children have needs and interests of their own and that the curriculum should present material for the solution of their immediate problems. One group holds that education is a preparation for life, the other holds that it *is* life.

The conflict would be partially resolved if it could be assumed that in the years of youth the individual would encounter all the types of problem of adult life. But no study of this has been made and partial examination seems to substantiate the claim that such would not be the case. Consequently, to construct a curriculum that would involve both of these ideas it is necessary to determine those objectives, interests, and needs which are highly important for children but are low in importance for adults. That there are many such is clear. Fairy stories, children's games, Mother Goose rhymes, and the choice of an occupation are a few items which among adults rank low in importance but in school curricula seem to be of great importance.

The discovery of these objectives of children which would supplement the important items of adult daily life is a necessary part of the study of the nature of children.

And not only do children's interests need to be determined. Their ability affects profoundly the sequence units in the curriculum. For it is evident that the units cannot be economically presented before the requisite maturity and experience for their assimilation has been secured. If educational psychology could give practical assistance in these respects the labor of curriculum construction would be greatly lightened.

The Trial and Error Method. In the absence of assistance from this source the division of the curriculum into grade units has to be accomplished by the trial and error method. The framer of curricula must in the light of his own experience and that of the classroom teacher make what in the combined judgment of both seems to be the best arrangement and he must then have it taught to children to see what modification, addition, or elimination must be made.

This, in itself, is a complicated and laborious process for which an adequate technique has not yet been developed. It involves both the

training of observers and the construction of standards. It requires extended verification by large numbers of teachers and classes and patient willingness to take almost unlimited pains. And all of this is foreign to the training and habits of school people who are accustomed to settle such matters in an offhand way upon the basis of unscientifically determined personal opinion. But, again, it is no less than any worker in the sciences accepts as a matter of course. The scientist cannot be parsimonious of either time or ease. And curriculum construction cannot be scientific until the same amount of patience, intelligence, and drudgery is expended upon it.

As a supplement to the consideration of the factors which have affected the failure to reconstruct the curriculum with greater dispatch I should like to present certain directions for curriculum construction which are for the most part implied in the foregoing discussion, and to describe some of the scientific investigations that have been made in the field. But before doing so it is necessary to mention two factors which operate in the construction of all school curricula. One of these is the element of time. The other is the element of curiosity and intellectual interest.

Intellectual Interest as a Factor in Curriculum Making. The claim is made for many school subjects, particularly in the high school, that their justification as a subject of study is intellectual interest. It is pointed out that the men who work in them are not actuated by any other life objective than that of satisfying intellectual curiosity and that the fostering of this capacity of the human mind should be one of the duties of the schools.

While this would seem at first glance to introduce a new factor into curriculum construction, a more extended consideration will reveal the fact that in the list of major life objectives the satisfaction of curiosity would or might be listed and that it would be treated in the same manner as the others. If it were found to be one of the most important it would have to be given a prominent place in the curriculum, but if it came low in the list it would receive less attention, or if it were found that it was not peculiar to one group of subjects but was an objective which was connected with every subject it would receive a different sort of treatment.

The Time Factor in Curriculum Construction. The element of time is extremely important in the determination of the curriculum. For after all of life's objectives are analyzed and the subject-matter which has been developed by man to reach them were collected we should have a mass so great that acquisition would be physically impossible for any individual no matter how many centuries he lived. He could not even keep pace with new knowledge.

Opposed to these centuries is eight or twelve years of school life during which some of this material must be mastered. Evidently, therefore, there must be a right selection of curriculum items. High indeed is the honor of nomination to this academy of the immortals. Not lightly in an afternoon faculty meeting can the selection be made. Nor

can we say that any members of this academy have vested rights in their position of honor. A new appraisal is necessary and each entrant must show his passport of distinguished service to humanity.

This paucity of years necessitates, therefore, an evaluation of the objectives of daily living upon the basis of importance and their arrangement in serial order from the most valuable to the least important. This list should, in turn, be modified by a similar evaluation of a supplementary list of the objectives in the daily living of children. And when these are formed the procedure to follow should be the determination of the number of these that can be taught in eight or twelve years with due regard to the interest and ability of the children.

In the construction of a school curriculum there will be eliminated those objectives which can be administered by agencies other than the school.

Directions for Curriculum Construction. These factors and their implications give seven rules for the construction of a curriculum.

1. First, study the life of man in the social setting and determine the ultimate objectives toward which he is striving.

2. Analyze these objectives and continue the analysis until working objectives are obtained.

3. Arrange these in the order of importance.

4. Raise to positions of high rank in this list those objectives which are high in value for children but low in value for adults.

5. Determine the number of the most important objectives which can be mastered in the time allotted to school after deducting those which can be learned outside of school.

6. Collect the best practice of the race in teaching these objectives.

7. Arrange them in their proper order of sequence according to the psychological nature of children.

This is a stupendous task for which few methods have been used and less have been perfected. But while the program is vast, it is no more so than that of any of the great fields of knowledge, and if the technique for performing the work can be worked out in a generation we shall be more fortunate than any of the sciences have been. The sole *desideratum* of final success is conscientious and intelligent effort directed upon problems which are of workable dimensions. For it is not necessary for any one individual to attack the whole problem. Science growing particle by particle has developed into an imposing structure, and the development of adequate methods of curriculum construction should be content to grow in the same inconspicuous manner.

Types of Curriculum Study. Fortunately for the progress of further investigation a variety of studies, crude and partial, have been made which are suggestive of the tendencies of investigators.

Spelling. The spelling curriculum has been studied more extensively than any other. The assumption has been made that spelling is of sufficient importance to be taught in school and the variety of studies has been determined by differences in the objectives which it is assumed to control. These studies are partial because an immediate objective is

assumed rather than one deduced by analysis from the major life objectives. They are partial, also, because they are not widely extensive in the material analyzed. But they are an improvement over the method used in obtaining our present spelling lists.

Four objectives have been assumed—that spelling is useful for reading, for adult written correspondence, for children's writing, and for correcting the errors of children. While these sound similar they produce lists which differ in items and in the relative importance of words as determined by frequency of use. This is apparent when it is known that the list of words useful in reading is obtained by an examination of newspapers, the Bible, and other authors; that spelling useful to children in writing requires the history of words in children's school compositions; that a vocabulary used by adults in written correspondence is found by the examination of business and personal letters; and that the errors of children require the listing of, not all words used by children, but those in which spelling errors occur.

The study of the spelling curriculum has proceeded one step farther and an attempt has been made to arrange the words by grades. One method used was to find the grade in which the most frequently used words first appeared. By another and more comprehensive method the words from several lists were assembled according to common occurrence in several of the lists and were then arranged tentatively by grades and later were taught to children for re-arrangement and elimination. In short, the spelling curriculum has been carried thru steps six and seven with four assumptions concerning the first five steps. These assumptions are expressed in the four objectives mentioned above.

Difficulties as a Curriculum Basis. In a recent publication Bobbitt advocates the construction of the curriculum on "the shortcomings of individuals after they have had all that can be given by undirected training". This method has been used quite extensively in recent investigations. One study assumes the function of grammar teaching in the grades to be the correction of grammatical errors of speech, and lists the grammatical errors of children as found by teachers and in the writings of children, classifies them according to rules broken, and determines the rules and definitions of grammar necessary for this purpose. The study in spelling mentioned above lists the errors in order of frequency of use, and this study lists all the errors made by children in a high school for a year. In manual arts it was assumed that the function was to produce ability to use carpenters' tools in repair work. The jobs around the home which needed attention were listed and the most common were ascertained. These consisted of such projects as putting up clotheslines, tightening screws in furniture, papering rooms, reseating chairs, stopping rat holes, mending windows, and adjusting doors. On this basis the tool processes necessary for performing these items were ascertained and evaluated on the basis of frequency. And in physical education studies of high school students have been made to discover their physical defects and corrections for these have been collected.

Cumulative Opinion as a Method. Cumulative opinion has been used as a method of determining the relative importance of objectives. Business men have been appealed to to determine what items in arithmetic should be eliminated or emphasized. Bankers have been consulted to discover the information of most use to depositors. And experts have been asked to determine the functions of American history.

Studies have also been made to determine what items of one school subject are useful in other school subjects. Studies of historical references found in civics textbooks, of the mathematics needed in the study of texts in chemistry and physics, and of the scientific vocabulary needed to read farmers' bulletins intelligently have been made.

The Analysis of Projects. Another type of study is the analysis of industrial processes. Analysis of the steps in farming projects has been made extensively by Stimson in Massachusetts. Similar studies have been made of the trades by the Federal Vocational Board, by the Emergency Fleet Corporation, and by the Federal Department of Labor. The analysis carried on by these agencies is most significant for all subjects of the curriculum. Their results are another example of the influence of new and critical situations upon progress. For, because they were working in new fields with poorly trained teachers and because a great responsibility was placed upon them, they worked out new methods which under ordinary conditions might not have developed for many years.

The Analysis of Major Objectives. Finally, some studies have been made in civics with the object of determining what citizenship means in terms of civil activities to be carried on in specific communities. But so far as the writer knows these are the only studies which have been made in connection with the first of the steps in curriculum construction mentioned above.

This partial resumé of the studies which have been undertaken is presented for the purpose of calling attention to the types of investigation that need to be carried on before the curriculum can claim to be scientifically constructed. The methods are crude and the studies incomplete. They lack the polish of chemical or physical investigation but they are worthy of attention because they have been made conscientiously and because the only method of obtaining better methods is to use and improve those already in use. Every science must pass thru a pioneering stage such as this, and twenty years from now the technique of curriculum construction will display the polish and accuracy which it does not now possess.

Next Steps in Educational Measurements

WALTER S. MONROE, Director of the Bureau of Coöperative Research,
Indiana University

THAT phase of educational research which is known as educational measurement has been a very conspicuous factor in the progress of the decade which is just closing. When sufficient time has elapsed to give a proper perspective, the decade from 1910 to 1920 will doubtless stand out as marking the beginning of a distinct epoch in the development of our educational thinking.

The beginning of educational measurement in this country dates from Rice's pioneer studies. The first of these, a study of spelling, was reported at the Cincinnati meeting of the Department of Superintendence in February, 1897. His underlying principle that the primary result of teaching spelling was the engendering of ability to spell and that one could measure the results of teaching spelling by finding out how well the pupils could spell, not only failed to receive the approval of his audience but precipitated an extreme and violent attack upon the American inventor of educational tests. This criticism, not only in the discussion following the reading of the report but in the educational press as well, was directed not at the results of the investigation nor the method of the investigation but at the underlying principle. Such criticism is evidence that a quarter of a century ago American educators believed that the purpose in teaching school subjects, such as spelling, was not to teach the children to spell but to develop their minds, and they were violently opposed to the suggestion that the results of teaching could be measured.

The work of Rice, which included studies in arithmetic and language as well as in spelling, did not affect school practice directly, but, what is more important, it profoundly impressed other scientific students of education, notably Professor E. L. Thorndike. In 1908, C. W. Stone, a student under Thorndike, published his arithmetic test, and at the Boston meeting of the American Association for the Advancement of Science in December, 1909, Thorndike presented his scale for handwriting. These two events mark the real beginning of scientific measurement of educational products in this country. Beginning with the arithmetic tests devised by Curtis, who was inspired by the work of Stone and Thorndike, the number of tests has multiplied with marvelous rapidity. Today any determination of the total number of tests is almost certain to be out of date before it can be published. A year ago it was well above one hundred. Not only has the number of tests multiplied but a large number of copies of several tests are being used every year. For at least two series of tests the total number published

has passed beyond the million mark and others will reach this mark at an early date.

The rapid growth of the science of education is also indicated by the establishment of bureaus or departments in colleges and universities for the purpose of carrying on educational research and for stimulating the application of the results of research in the public schools. For the most part the work of these bureaus has been confined to the field of educational measurements. In addition, a rapidly increasing number of school systems are making distinct provisions for the systematic use of educational tests by the establishment of a department having for its special function the supervision of the use of educational tests within the school system.

Before such an audience as this it is not necessary to recall in detail the activities of these bureaus or departments. In general, their activities fall into two groups, service and research, altho for the most part the research problems studied have been chosen with reference to making possible increased service. State bureaus established in colleges and universities have rendered service to the public schools of the state; *first*, by making educational tests easily accessible to the superintendents and teachers; *second*, by providing necessary accessories such as directions, class record sheets, and the like; *third*, by instructing teachers and superintendents in the giving of the tests, in tabulating the scores of the pupils, and in calculating the median or average scores; and *fourth*, by providing a clearing-house where the reports from the several cities were assembled and arranged in a form suitable for making comparisons not only among the cities but also between the cities and standard scores or scores from other states. Research has been carried on in the determination of the relation between the results of teaching and the course of study, in the determination of the efficacy of certain methods of teaching, and in some instances in the derivation of new tests. The activities of bureaus organized within a school system have been similar. They have rendered service by systematizing the use of educational tests within the city, by gathering together the results for the city as a whole, and by assisting in modifying instruction upon the basis of the results secured.

These outward manifestations are indicative of a fundamental change in our educational thinking which is now taking place. Within a brief quarter of a century, to a large extent within a decade, we have replaced the thesis that the result of the process of education is a general development of the mind which is not measurable by the thesis that the results of education include as fundamental elements specific habits and skills which may be measured with precision and also that many of the other outcomes are capable of objective measurement. The incorporation of this change, which is now generally accepted, in the educational thinking of superintendents and teachers will mark a more significant accomplishment than the multiplication of the educational tests and the establishment of organizations to stimulate their systematic use. It will mean that we are thinking of education as resulting in definite outcomes which we may definitely work for. Our purpose will be the engendering

of these outcomes rather than the mere teaching of certain subject-matter. Our courses of study will be written in terms of definite objective aims instead of general qualitative statements. For example, the teacher of handwriting in the fifth grade will be told that she is expected to have her pupils able to write at the rate of sixty letters per minute and with a quality of sixty-five on the Ayres Scale, instead of being told that she is expected to have her pupils write rapidly and legibly. In silent reading the aims will be expressed in terms of words read per minute and definite degrees of comprehension. In the operation of arithmetic the rate and accuracy at which pupils are expected to be able to do certain types of examples will be specified.

As one reflects upon the accomplishments in the field of educational measurements during the past decade there comes the conviction that we are on the eve of significant developments and that state bureaus of educational research will have a prominent part in them. Much of the service which we have been rendering will no longer be needed. Soon it will not be necessary to instruct superintendents and teachers in the giving of educational tests and in the tabulation of results nor to urge their use. By means of the work which they are now doing they will have acquired this technique. Furthermore, they will have come to recognize the value of using these tests as instruments for making their efforts more effective. In the field of research, the problems studied have for the most part concerned the variability of pupils, classes, and cities. The conditions revealed have been corroborated so often that we may regard them as established facts, and altho comparisons of scores will continue to have a value, standards for widely used tests are now generally available for anyone who wishes to compare his scores with the standards. On the other hand, new problems are arising with which we must deal and it is to certain of these that I desire to direct your attention.

1. The Evaluation of Existing Educational Tests. The number of available tests has increased until we now have several for most of the subjects in the elementary school. Some of these tests have been constructed with care and upon the basis of sound procedure. Others have been carelessly constructed. In certain cases simplicity of form and convenience in use have been prime considerations; in the construction of others these factors appear to have been given no weight. Some tests are accompanied by a multitude of accessory directions, record blanks, and graph sheets. Other tests are not even supplied with directions for use. Some tests are designed to yield average or general measurements while others have been designed to yield specific or diagnostic measurements. Superintendents and teachers are already beginning to ask what is the best test in reading, in arithmetic, or in some other subject. "Is Test A better than Test B?" At the present time a rational choice of a test or of a group of tests in such a subject as silent reading or arithmetic is difficult if not impossible because there is not available the necessary information concerning the function, validity, reliability, precision, comparability, and similar features. Much information on these points for certain tests is scattered thru our educational

literature but nowhere is there to be found a complete account for any one test. There are a great many things about widely used tests which would be helpful to know but which yet remain to be determined. For example, the probable error or the coefficient of reliability is known only for a very limited number of tests and it is obviously necessary that we know the degree of reliability of our measurements before we can be certain in our use of them as the basis for planning instruction or in determining the relative merits of different methods of instruction. We need, therefore, a critical evaluation of our educational tests. In part this can be done by bringing together and organizing the facts which are now scattered thru our educational literature, but much original research will be required. In my judgment this is one of the most important research problems with which a bureau of educational research can deal within the next few years.

In order to make a rational recommendation of educational tests it is necessary that we have for a test the answers to the following questions:

1. What ability or group of abilities was the test designed to measure?
2. Was the test designed to yield general or diagnostic measures?
3. What are the significant characteristics or dimensions of the ability or group of abilities which the test was designed to measure?
4. What characteristics or dimensions does the test measure?
5. Does the test measure the ability or group of abilities it was designed to measure? Does it measure this completely? Is the pupil's performance on the test determined by this ability or group of abilities or is it determined to a variable extent by other factors such as his rate of writing, his language ability, his general intelligence, or the manner of the examiner?
6. Have the exercises been designed so that they cover or are representative of the field of ability which the test is designed to measure?
7. Will the acquaintance with this type of test influence the pupil's performance? If so, how much?
8. To what extent is it possible for pupils to make special preparation for the test and thereby make scores which are not valid indices of their abilities?
9. In comparing the scores of individual pupils, schools, or cities, how large must the differences be in order that these differences be significant? What is the variation in terms of probable error or average deviation of individual and of group measurements? What is the coefficient of reliability?
10. When it is desired to give a test a second time in order to determine the progress which pupils have made what allowance must be made in comparing the two sets of scores?
11. How much time is required from pupils and from the teacher for the giving of the test and tabulating the results?

The answers to some of these questions should have greater weight than others in determining our recommendation. For example, if a test is to be used by teachers it must not require an unreasonable

amount of their time. Also the test must be accompanied by the necessary accessories.

When tests are subjected to a critical evaluation such as has been suggested we will undoubtedly find that many, possibly all, of our educational tests are relatively crude and that there is a distinct need for more refined measuring instruments. For example, up to the present time relatively little attention has been given to the analyses of the subject-matter fields in which tests have been constructed. We do not have an analysis of the subject-matter field of silent reading altho we have many silent reading tests. The need for such analyses is more obvious in such a subject as spelling and as a result we have had numerous determinations of the words which pupils should be asked to learn to spell and consequently which should be used in constructing a spelling test. Most of the analyses which have been made are best known in connection with the determination of the minimal essentials or scientific curriculum making but they are also fundamental in constructing educational tests. In the process of refining our present tests or in constructing new ones it will be necessary to invade this field.

The mental processes involved in the different school subjects must also be analyzed so that we can know just what the essential characteristics or dimensions are of the abilities which we are attempting to measure. In the operations of arithmetic we have found that we have to measure not one ability but many abilities which are to a certain degree independent of each other, and that the essential characteristics of these abilities are the level of difficulty, the rate of functioning, and the accuracy of the functioning. This condition places certain definite requirements upon tests in this field. In other subjects, similar analyses have not been carried as far. In silent reading the indications are that there are a number of different kinds of reading ability but the analysis is incomplete. We have at least six different silent reading tests which have been widely used and it appears that no two of these measure the same silent reading ability. This condition indicates the need for a scientific determination of the relative importance of these silent reading abilities and whether these are all of the important silent reading abilities. Until such determinations have been made, selection of a silent reading test will be largely a matter of personal opinion.

These are two lines of allied research which must be entered before we can have properly refined instruments for measuring the results of teaching.

We should realize that educational measurement is not an isolated topic in the field of education but that it has many vital connections. One cannot go far into this topic without forming many contacts with other topics, and students in other fields of education are constantly making contacts with educational measurements. In making this point I am not advocating an imperialistic program of expansion for the field of educational measurements, but I am simply indicating the dependence of this department upon other lines of research. Workers in the field of educational measurements must acquaint themselves with the results

in these allied fields and probably even enter them in order to secure fundamental information.

2. **More Refined and Complete Interpretation of the Results.** We must also refine and enlarge our interpretation of the results of educational measurements. The meaning of scores is limited and depends upon factors other than the differences between them and the standards. Not infrequently misleading conclusions have been reached because certain potent factors were neglected or because the scope and function of the test were neglected. For the most part up to the present time our interpretation has been confined to the comparisons of city scores or school scores. There is a distinct need for extending our interpretation to include the scores of individual pupils.

Any test measures the results of instruction only over a very limited field and we must constantly bear in mind these limits in interpreting the results. We are accustomed to speak of a test in the field of arithmetic as an *arithmetic test* and thereby imply that it measures all of the results of teaching arithmetic. Unless it consists of a very elaborate series of tests it is impossible for a single test to cover this entire field. The most widely used arithmetic test is the Curtis Standard Research Test, Series B. It covers only the field of the operations with integers and it covers this only in a limited way. It is, therefore, misleading to think of this test as measuring the results of instruction in arithmetic. It measures only certain results of instruction. There are several types of examples in the operations with integers besides the four operations with fractions, both common and decimal, and the whole field of problem solving concerning which this test gives us no information. Therefore, in interpreting the scores of this test it is necessary to bear in mind the limits within which it measures the results of instruction. To conclude, because a school makes scores up to or above standard on this test, that the results of instruction in arithmetic are satisfactory is to reach a conclusion not based upon fact and which in some cases will be erroneous. Of course one may take the position that even if this test does not cover the entire field of arithmetic, the scores are representative of the entire field. This assumption may be justified when the test is given for the first time but where it is used regularly the outcomes measured by it are likely to receive special emphasis in teaching and this has been found to be the case in some cities.

A test in the field of reading is generally spoken of as a reading test, which implies that it measures all of the results of instruction in reading. Reading appears to be an exceedingly complex field and any one test probably covers only a small portion of that field. It is, therefore, misleading to consider the scores of any one test as representing all of the outcomes of the teaching of reading. We need to analyze our reading tests so that we may know the limits of the field of ability which each measures. Until we do this a large element of uncertainty and possibly error will be involved in our use of reading tests. What is true of arithmetic and reading is also true of the other subjects for which tests are available.

There is a need for another type of refinement in our interpretation of class and city scores. There is a distinct tendency to rate a school system or a class which has median scores higher than other similar groups as being more efficient. We have not taken into consideration in this interpretation the fact that both school and city scores depend upon a number of factors other than the quality of the instruction which the pupils have received. Obviously the results of instruction depend upon (1) the time allotted to the subject by the course of study, (2) the length of the school term, (3) the regularity of attendance, and (4) the promotion rate. Neglecting the promotion rate when comparing two cities is to place a premium upon a low rate of promotion. For it is obvious that the scores of any grade can be increased by holding in their grade pupils who make relatively high scores and who, therefore, should be promoted to the next grade, and also by holding in the grade below pupils who make low scores and who because of their standing in other subjects or because of their age should be promoted. The scores of a class or city also depend upon the elimination of pupils and provisions for segregation of backward pupils. Some of these factors are doubtless more influential than others, but before we can make a reliable interpretation of a class or city, it is necessary that we take all of them into account. The development of the technique for doing this constitutes one of the research problems with which bureaus of educational research must deal in the future.

The interpretation of scores must be extended to individual pupils or at least to groups of pupils within the class. Mass interpretation is crude. It tends to eliminate the characteristic differences of pupils and to reduce all to a common level. That pupils differ widely in their achievements is demonstrated every time a test is given to a representative group. They also differ in their needs and in respect to the kind of instruction to which they will respond. A certain type of instruction will be helpful to one pupil or to a group of pupils while to others it will be of little or no assistance.

In extending interpretation to individual or group scores it is necessary to keep in mind that pupils having the same scores do not necessarily have the same instructional needs. For example, two pupils may make low scores for entirely different reasons. It will, therefore, be necessary frequently to supplement the scores by additional information. Sometimes this can be secured by analyzing the pupil's performances which are recorded on the test paper. Occasionally, it will be necessary to subject the pupil to an individual examination in order to study more fully his mental processes. It thus becomes necessary that we know the significant characteristics of various kinds of performances and the meaning of each. We must be able to translate the errors of pupils and also their creditable performances into their needs for future instruction. The educator needs to do what the physician has done. When a physician is called upon to prescribe treatment he examines his patient to determine his medicinal needs. He does this by noting certain characteristics such as pulse rate, temperature, appearance of the eyes and tongue,

etc. A physician who knows the meaning of the facts he finds easily translates them into the needs of his patient.

For the most part this is a virgin field, altho it gives great promise of profitable returns for the time and effort which may be invested in it. We need to collect and study the errors which pupils make. All errors are not alike. If one hundred misspellings of a word are examined it will be found that a few ways of misspelling the word have been repeated many times while others occur only once. A misspelling which recurs twenty times has a meaning different from one which occurs only once. What is true of errors in spelling is true of errors in other subjects. We need to study the performances of pupils so that we may know what they mean.

3. **Modifying Instruction on the Basis of Educational Measurements.** Many of those who have contributed to the development of educational measurements have stopped their work with the measurement of the results of teaching. Some have actually advocated that it was not wise to go farther and that the improvement of instruction should be left to workers in the field of methods. However commendable this may be from the standpoint of professional courtesy, it has not proven satisfactory in practice. There has been much waste in the use of educational tests because emphasis has not been placed upon the use of the results in the improvement of instruction. Frequently the work has stopped with the tabulation of the scores of the tests, and both teachers and superintendents have expressed dissatisfaction with the tests because they were not able to detect any improvement in their pupils. Not long ago I inquired of a superintendent in a city of about 50,000 population concerning his use of educational tests. He remarked that they had used the tests, in fact they made it a point to keep up-to-date and to give new tests in at least one or two schools. He added what was to me a very significant statement. He said "We gave the Courtis Standard Research Test in Arithmetic last year but the teachers failed to find that the pupils did any better in their arithmetic work after the tests were given than they did before."

It is obvious that this superintendent did not realize that educational tests are not teaching devices but instruments which a teacher may use to secure accurate information concerning the results of her teaching and that whatever improvement in instruction may take place will come not from the mere giving of the tests but from the use of which is made of the results of the tests in the planning of future instruction.

The planning of instruction on the basis of scores secured from a test so as to correct the defects revealed and provide each pupil with that instruction which will enable him to make the most rapid growth of which he is capable is more complicated than simply applying special or general methods. It is first necessary to interpret the scores in terms of pupil needs, and this requires an understanding of both the test which was used and the nature of the mental processes involved in the subject. After the scores are interpreted the modification of instruction involves the application of general principles of method to particular

cases. The position is now generally being taken by measurement men that the use of tests should always result in the improvement of instruction and that they must carry their work far enough so that the classroom teacher will know how to make use of the information which the test provides. Thus it becomes necessary that as an integral part of its work, a bureau of educational research should study the performances of pupils, particularly their errors, and give attention to collecting and devising methods of instruction which will correct these errors.

The carrying on of educational research along the lines which I just indicated, as well as in others, represents only one phase of the work of a bureau of educational research. Research in the field of education is valuable for its own sake and there will always be need for workers who are interested in contributing to the accumulation of human knowledge, but if there is not to be great waste there must also be those who contribute to the application of the results of research in the public schools. Therefore, it is vital that state bureaus should recognize that the stimulation of the application of the results of educational research in the public schools is one of their important functions. In fact, it is my judgment that for the immediate future, service to the public schools of the state should be made paramount in activities of state bureaus of educational research and that the research undertaken should be chosen with a view of increasing this service. This will not result in limiting the research activities for as I have indicated there are numerous vital problems which must be studied before we can render satisfactory service.

In bringing to your attention these "next steps" I have no intention of minimizing the use which has been made of educational tests in the past or any use which you may now be making. Some or all of these steps may not be new to some of you. Much progress has been made but the final goal has not yet been reached. No matter how creditable our past or present achievements may be, we must not be content to stop or even slacken in our efforts to achieve greater things. In the field of educational measurements this is not merely a commonplace generalization. As one reflects upon the development in this field during the past few years and upon the present activity it becomes clear that we have reached or are rapidly approaching a distinct mile-post. We are devising tests, standardizing them, interpreting the scores, and using the results in much the same way as we were three or four years ago. Our tests are not greatly different in kind and have not been improved except in relatively few instances. It, therefore, is pertinent to ask, What are the next steps?

I have outlined to you certain important next steps and urge you to consider them in connection with the basic question, What are the next steps? The evaluation of existing tests so that rational choices will be possible is primarily a research problem but it is one in which your bureau of educational research will need your coöperation. The refinement and extension of the interpretation will require a greater investment of time and energy in the study of scores and until a technique for doing it has been worked out, it will be difficult. But we should always

bear in mind that only as we know what scores mean are we able to use them and that to attach a wrong meaning to them is disastrous. The value of educational tests is realized in the use that is made of the resulting scores. When the use of an educational test stops with the interpretation of the scores and there is no modification of instruction, the time and money invested is largely wasted. In order that the use of the test may be profitable, the scores must be interpreted in terms of pupil needs and instruction planned to meet these needs. And, finally, we must engender in teachers the realization that educational tests are instruments which may be used by them as well as by superintendents and principals, and that their primary function is to secure information which will be helpful in planning future instruction. I look forward to the day, and I do not believe it is very far in the future, when educational tests will be considered necessary items of schoolroom equipment just as we now consider charts, blackboards, reference books, supplementary readers, and the like necessary for the most effective instruction, and to know how to use educational tests wisely will be considered an essential factor in the teacher's professional preparation.

Educational Service in Iowa

ERNEST J. ASHBAUGH, Director of the Bureau of Educational Service,
Extension Division, University of Iowa

THE Bureau of Educational Service at the State University of Iowa has been organized almost five years. During this time it has been under the present director and followed a rather regular and somewhat uneventful career. There were no broken paths to follow at the time of organization and considerable time had to be spent in feeling the way and gradually learning the fields in which there were needs and opportunities to work. Being new to the state, the superintendents, principals, teachers, and school boards were not at all sure whether the new man and the new bureau could be trusted with the inside workings of their educational systems or not.

The earliest efforts were made in the field of child accounting and teachers' marks but the response revealed the fact, which perhaps should have been foreseen, that somewhat simpler forms of data must be gathered first. However, the data that were secured resulted in revealing some things very much worth while from the standpoint of the development of the work of the bureau.

1. It was found that in spite of the fact that the schools were reporting a very high percentage of attendance, seldom below 90 per cent and more often between 95 per cent and 100 per cent, that on the basis of fifteen school systems and over 13,000 children about 20 per cent were out more than 10 per cent of the time. Ten per cent were out more than 20 per cent of the time and 6.6 per cent were out more than 30 per cent of the time. And this was in a year in which we had no epidemic. These data, showing the superintendents an unexpected amount of absence, were valuable in two ways. They caused a much more searching attitude on the part of the better superintendents with respect to the whole question of absence. A number instituted a closer check on the cause of the absence and some very interesting results were thereby revealed. At least two men presented the data which they gathered in this connection before their Parent-Teacher's Associations which data showed conclusively that a very large portion of the absence could have been prevented by the exercise of ordinary care on the part of the home. The other result, which from the standpoint of the bureau was perhaps even more pleasing, was that it gave a decided impetus to the belief that the Bureau of Educational Service of the University was really a bureau of service and could assist superintendents in discovering things which it was worth their time to know.

2. The semester marks given by the teachers to these pupils were checked in terms of these absences and the evidence of a positive correla-

tion between absence and poor grades was most convincing. Contrary to general belief, however, absence seemed to be no more destructive to marks in the upper grades than in the lower. For example: 5.3 per cent of the children who had perfect attendance in the first grade received failing grades, and 30.2 per cent of the children who were absent more than 30 per cent of the time failed. In the fourth grade 6.5 per cent of the children having perfect attendance failed, while 27.8 per cent of the children who were absent more than 30 per cent of the time failed. In the seventh grade, 3.7 per cent of the children who had perfect attendance failed, and 25.4 per cent of the children who were absent more than 30 per cent of the time failed. The average for all grades showed 5 per cent of those having perfect attendance received failing grades, while 23.7 per cent of those who were absent more than 30 per cent of the time failed.

3. The distribution of teachers' marks showed just the type of thing that these investigations have always shown, namely, a very wide range of variation. But in some cases the revelation was a surprise to the superintendent and resulted in closer supervision of the pupils' grades. For example, one city gave grades of 95 per cent or above to 24.5 per cent of its pupils while another gave the same high grade to but .7 per cent of its pupils. The first city gave failing grades to 3.4 per cent of its pupils while the second gave failing grades to 18.6 per cent. The first superintendent wondered whether he had such very superior teaching as to warrant the high percentage of high grades while the second wondered whether the teaching in his system was poor or the standards of attainment expected were too high. A careful analysis of the situation resulted in some very desirable adjustment.

The bureau next took up the work of stimulating the use of standard tests and the state-wide survey of single problems. Arrangements were made for the bureau to be a distributing center for the various standard tests on the market. This work began in a small way and has been extended as the interest in the work has grown over the state. No special inducement was offered in any way to the superintendents to take up the work. The interest grew as increasing numbers of superintendents, principals, and teachers attended summer school and took courses in tests and measurements. In county institutes and at teachers' associations the value of such testing was demonstrated and in a few cases the director went into individual cities and supervised the giving of the tests.

While more or less testing has been done in this manner in more than 150 different cities and towns in the state, the results in some cases have not been as satisfactory as could be wished. Sometimes the teacher or superintendent, knowing only a little about the measurement movement, has sent in for a few tests, given them, computed the medians or averages, and considered the thing done. He has not realized that only the first step has been taken and that the value of the testing, if it is to have any real value, must come from an analysis of the results, a diagnosis of the causes lying back of these results, and the application of the proper measures to insure the highest type of result. Too often

low results have been summarily dismissed with one of the following inferences: (1) The teacher is poor; I'll get another for her place next year. (2) What is needed is more time; we'll just spend more time on that subject and that will fix it. (3) Well, I don't think that test amounts to much. You can't measure education or the influence of a teacher like corn or hogs or lumber anyway. You know how fallacious such judgments are, but *we* have had some school people who have made them and acted upon them.

Because we have felt in Iowa that the measurement work has not accomplished all that it ought to accomplish, we are planning this coming year to extend to the schools doing testing a very definite offer to assist them in interpretation, diagnosis, and treatment. We shall not force this upon anyone. We shall make it clear that this type of service is open to them and that we are eager to be of greater assistance than ever before. We shall probably spend some time and energy in demonstrating the worth of this type of work in order to stimulate the requests for it.

State-wide work has been done in writing, Courtis arithmetic, and spelling. This work has been helpful in at least two respects: It has given to the superintendents of the state comparable data from other schools within their own state against which to check the status of their own system. We find that our superintendents would much rather check their own school results against results of neighboring schools, or against Iowa as a whole, than against standards which represent the entire United States. In neither of these subjects has the work gone to the extent to which it might be pushed. In all cases the only phase that has been taken up has been that of status. The question of method, of time, of experience and training of the teachers, of the extent of supervision and continuity of the administration have not been checked out carefully. Work in other subjects has been more or less fragmentary, but fall and spring testing under carefully controlled conditions will be attempted this coming year.

During the past two years another type of work has been growing very rapidly. In fact, this spring the calls could not be taken care of by the force available. This is in making surveys and giving advice to school boards relative to new buildings and grounds. Even during the war, some building was going on, but with the signing of the armistice a large number of communities which had been holding up any plans for the remodeling and enlargement of their school plant began to make preparation for this work. In consequence, a large number of calls have come in for this type of service. School boards want to know especially concerning the site, and the type of buildings to be built.

Where should a school building be built? There are many factors to be considered as you very well know. And yet I have been surprised again and again how often school board members have shown me by their statements that some of these factors have been left out of their consideration. That it should be located conveniently for the pupils to be accommodated is plain enough, but we must remember that buildings are not built just for today and the center of present school population

is of less importance than the center of population as it will be during the major part of the next twenty or thirty years. I am thinking now of a high school building in one of our Iowa cities which is now an old building and must be replaced soon. It is very near the center of the present high school population. It is upon the site of the original high school building which the city erected nearly three-quarters of a century ago. Because of this fact, there is a very strong sentiment in the city that the new building must be erected on the same site, in spite of the fact that the city must grow almost entirely in one direction. An analysis of the registration of voters for the last twenty years, of the location of the cable lines put in by the telephone companies, and of the extensions of the water and gas lines all point to this one conclusion, that the city must make its growth in this one direction. I am glad to say that there is a growing tendency in our state for school boards to ask the bureau of educational service for assistance of this kind in analyzing their problems and in acting upon the recommendations after they are made.

Perhaps another illustration will make the matter even clearer: A few weeks ago the director was asked to visit one of our small cities, a place of about six thousand, to answer the question of type of building and site. The city was a rather typical agricultural center lying on a level plain. The possibilities of extending the city's growth were practically as good in one direction as in another. There were no hills or valleys or streams to interfere. Two railroads passed thru the city at almost right angles with each other. Two manufacturing plants were being solicited by the commercial club for establishment in the community. On the basis of a receptive attitude on the part of these industries, real estate men were endeavoring to impress upon the school board the necessity for building buildings in the additions in which these industries would locate, if . . . To complicate the matter somewhat further, the school board now owns a plot of ground within half a block of one of the railway stations, and on one of the main streets. Because a school building erected upon that lot would be passed by a majority of the people entering and leaving the city by rail, and because the street leads out on to one of the main traveled highways, the majority of people entering and leaving this city by auto would see the building, it was considered a highly desirable site since the building would advertise the city.

You will note that the items being considered were not at all items pertaining to the education of the children. Being in a purely advisory capacity I could not tell the school board that they must not erect their new building upon this site. The best I could do was to half humorously tell them, "I could not give my consent." The possibilities of their growth, and the present buildings which they have, pointed very specifically to one thing, viz: a junior high school building located rather centrally but away from the railroads and business districts. The arguments from the standpoint of the children arrived at this conclusion alone. The arguments from the standpoint of the adult population probably pointed in another direction. In reporting to the school

board I made it my special point to insist that this proposal was in behalf of the children.

The second question, the type of building to be built, is more complex but also even more interesting. We have been working on the general principle that school communities should make provisions for as wide and comprehensive an education for their people as they are financially able to do. We are inclined to insist that this is the best possible investment that a community can make. We are quite sure that no child who is mentally capable of profiting by an elementary and secondary school education should be prevented from securing it either because of economic conditions in the home or because of the failure of the school to make the necessary provisions in its physical plant or curriculum or both. We said any child mentally capable, and we are quite sure that not nearly all those mentally capable have been remaining in our schools thru the elementary and secondary grades. In fact, when the curricula are sufficiently differentiated to include a very much larger number of kinds of work than has been offered in all too many schools in the past, the mentally incapable will have become far fewer than our present number of failures might incline us to believe.

In recommending for any community the type of building that should be built, the following factors are carefully considered: (1) the educational needs of the community as evidenced by the education of *all* its people; the vocational activities in the community; the vocational activities of the state and nation; the vocational desires of the children and young people of the community and the fundamental requisites in general education for democracy and Americanization; (2) the need for physical and social education by *all* the people and the extent to which other agencies are caring for this need, especially with the adult population; (3) the present school plant and its possibilities in the way of remodeling and utilization in the big program of education to be undertaken by the community; (4) the financial status of the community with the possibility of building all at once, a type of building which will accommodate this extensive program or the necessity for building a part and later adding as the finances of the community will permit.

I think a little time may well be spent in making more concrete some of the things considered under each of these factors. In undertaking to determine the type of building that should be built, the first consideration is certainly the educational needs of the community. A very careful survey of the present education of all the people in the community is quite necessary. We have all too long been considering that our school plant was for the education of children between the ages of six and fourteen, or six and twenty, and have failed to realize that the community should make provision for the further education of *all* its people who may desire to take advantage of such.

Some of our larger cities have been doing work in evening courses, especially in manual and domestic arts, commercial lines, and English for foreigners. These are well and good and should be made much more available than at present. There is no reason, however, why the English-speaking population who might desire it should not have an oppor-

tunity, as a part of the educational program of the community, to further their education in cultural lines as well as industrial. In fact, the rapid development of the educational status of any community is dependent upon the possibility of making an appeal to all the people rather than only to a small group.

In making a survey to determine the education of a community, specific questions have been asked concerning the education of the parents in terms of their completion or non-completion of elementary, secondary, collegiate, and professional schools, and in terms of the age of leaving and the educational status at time of leaving. When the object has been clearly made known and the way prepared, but little opposition has been experienced in securing this information.

This type of information is valuable not only from the standpoint of an enlarged educational program, especially in the nature of new buildings, but is also valuable frequently to the classroom teacher in working with the children. The educational status of the parents frequently furnishes an index to the educational opportunities such as high-class books and magazines which are provided in the home.

When 50 per cent of the parents of our high school children and a still larger per cent of the parents of the children in the grades (data from two cities only) left school at the age of sixteen or younger, and with a schooling of eighth grade or less, there should be developed in our public school system an opportunity for extension work which would enable a larger number of these people to secure additional educational advantages. As we stand at the present time, we have developed an attitude that when one becomes an adult his school days are over, except for those who go into the professional lines. We are definitely endeavoring to create a different attitude, viz: one which says that thruout life the doors of our public schools are open and provision will be made for extension work to care for any type of education which a group in the community may desire and the cost of which is not prohibitive. To be specific, classes in domestic science and art, in home nursing and allied fields, could with profit be run in the school as extension activities in even the comparatively small communities. That there is a need for such is not questioned. The only thing necessary is an adequate provision in the schools and the development of a sentiment in the community in favor of taking advantage of this sort of thing.

It is quite easy to ascertain the number and types of vocational activities in any community. In a sense it brings in the industrial survey, but it does more than that in that it secures much needed information regarding not only the different types of activities within the community and the number of people engaged therein, but in terms of the opportunities for advancement and the educational demands which the vocation makes.

Realizing, however, that only a comparatively small proportion of our people remain for a generation in a single community, the vocational activities of the state and nation must also be considered in planning the big program. Only large cities can provide specific vocational training for many industries within their borders, but comparatively small places

can make provision in their educational program for acquainting their own people with many types of vocations and with the underlying principles which workers in these vocations need to know.

Our school curricula in cities of five to twenty-five thousand are pretty generally providing for three types of education. (1) They provide for the broad use of the tool subjects which are fundamental for all education and for daily communication and life activities. (2) Provision is made in the secondary schools for admission to college and this is well since a very respectable percentage of our high school pupils go to college. (3) The third line of education is in the commercial field. Practically all of our high schools now have a commercial course. It is unfortunately true, however, that the high school commercial department is caring for but a small portion of its legitimate field. It is making a strenuous endeavor to train bookkeepers, stenographers, and typists, but in a large percentage of cases has quite overlooked the fact that the biggest phase of the commercial field is selling. This phase of commercial work must be built up and provision must be made, not only in our course of study and in our teaching force, but in the building of our school plants of the future, to take care of this third activity in this department. It is hardly necessary to mention the fact that much more adequate provision for home-making and manual activities must be provided. Thus the vocational activities of the community, the vocational desires of the children, and the vocational possibilities of the state and nation must all be taken into consideration in the determination of the type of building.

Scarcely less important than the provision for a broad use of the tool subjects and even more important than the provision for vocational or pre-vocational education is the need for provision for physical and social education. We are just beginning to appreciate the need of these things. Even today many communities are not building their school buildings in such a way as to provide for physical education, nor appreciate the need for provision for social education. The demand for adequate auditorium facilities which will make possible the bringing together of different classes or of all the classes of the school into a common assembly, or for the bringing together of the youth and adults of the community as a social group, is just beginning to be recognized as it should be.

An athletic field of sufficient size is still almost a hopeless asking in our cities of above five thousand. The gymnasium is pretty well recognized, but largely for the use of the immediate school pupils alone. A recognition that the gymnasium should be an integral part of the provision for the needs of the adult, out-of-school population seems to be but vaguely understood. There is a great opportunity for the use of a gymnasium for adult groups in evenings, especially in the towns which do not have a Y.M.C.A. We are making a special effort to assist school boards in understanding this opportunity and consequently to plan for a school plant which will take care of this need.

In few communities is it either possible or desirable to scrap the present school plant and build entirely new. It is often much to be regretted that this is not possible, but the situation must be dealt with

as it exists. Consequently, the existing plant must be considered from the standpoint of its utilization in the development of a big education program. Frequently a shift in organization makes possible a more advantageous use of the plant. Again, relatively inexpensive additions or changes within the building or group of buildings may make possible a much wider use than heretofore.

Finally, of course, the financial condition of the community must be given utmost consideration. It is not only useless but a positive detriment to present to a community plans which are wholly beyond attainment because of financial limitations. This work must be helpful, it must appeal to the school board and to the public as safe and sane, and while it must have large vision, it must at the same time be within the financial reach of the community.

This work has been discussed at so much length because it is a prominent type of our work at the present time. We are definitely cultivating this because we believe that it is an opportunity for the bureau to influence education in the right direction. There is always a tendency for the local boards to see only the local situation. Furthermore, there is always a tendency for them to see the local situation in the light of their own early educational history or experience, rather than in terms of a vision for the future. This reflects nothing upon the members of the board. They are busy men, devoting the major portion of their time to their own business affairs. It is not to be expected that they should all keep step with the most advanced things in the educational field. We believe that it is our special duty and privilege to assist them in gaining this big viewpoint and show them how they may develop it in their own local communities.

During the past two years another type of work has been greatly developed, viz: an appreciation on the part of the school boards and superintendents of an evaluation of the school system by a survey at the beginning of an administration. As the superintendents are learning more and more the nature of our survey work, they are appreciating the value of an inventory or evaluation of the educational status of the new system into which they are going. They welcome it at this time because they are not responsible for the conditions which exist. They appreciate the value of having detailed knowledge of conditions under which and with which they are to work. They realize, also, that this is a background against which any future progress may be placed as an evidence of their efficiency. Of course they appreciate, also, at least some of them do, that failure to produce may also show up against this same background. Most superintendents, however, are quite confident of their ability to secure results and therefore welcome concrete material which will show the starting-point from which they have progressed. School boards are appreciating that it is only just to the new man to give him every possible bit of knowledge which can be secured concerning their schools. They realize that this makes for higher attainments, for greater progress.

I believe the best evidence that the work of the Bureau of Educational Service is being appreciated in Iowa is the increased extent to which the

superintendents are responding to the request for coöperation and to the growing number of requests from school boards for specific types of service. In general, the plan has been thruout the five years to extend a general invitation to all superintendents in the state in cities of five hundred and above, to contribute on all general problems which would be interesting in all sizes of cities. The requests have always been issued in the form of an invitation to coöperate in case the superintendent feels that the data gathered would be of use to him. Special emphasis is placed upon this point. We have endeavored to make the superintendents feel that the cordial relationship existing between them and the bureau would be maintained regardless of whether they contributed or not, and that the important consideration was for them to decide whether it was worth the time and energy necessary to secure the information.

During the first year not more than a 3 to 5 per cent response was secured from these invitations. During the last two years, response has been secured from more than 50 per cent. The last data called for secured a 50 per cent response in a week's time. During the first year but one request came from a school board for special service. During the present year, more requests have come in than it has been possible to care for. In fact, work is now dated which will cover a number of weeks during the coming year. This gives you a general indication of at least three types of work which the bureau has been endeavoring to do, and some idea of the reception which it is receiving from the school people of the state. It is presented here because we believe in the service and believe that you have a bureau here at the University which can do any of these types of work, as well as others which have not been mentioned, to the mutual advantage of the local community and of education in general.

Some Recent Developments in Spelling

ERNEST J. ASHBAUGH

PROBABLY no subject in the entire elementary school curriculum has received more attention during the last few years than spelling. Perhaps the chief reason for this is that it is unusually easy to check,—at least it appears so on the surface. Also the initial studies in measurement were made by Rice and Corman in this subject.

During this period of development, which has now extended over more than twenty years, the movement has followed three chief lines: (1) the selection of spelling lists; (2) the method of teaching and learning; (3) the development of scales and tests for measurement of status and progress.

Far more has been done in each of these lines than it will be possible to discuss during this period. I shall pass with briefest mention the earlier work in each line in order to confine myself mostly to the more recent developments.

Spelling books have been in use in our schools since the days of the famous Webster Blue Back, and during that time it has run the whole gamut of variation from books containing a very small list to books containing more than ten thousand words; from those which have no suggestions on the teaching or learning process to at least one boasting over six hundred separate suggestions.

The method of selecting the word lists has followed four general lines:

1. General English vocabulary such as is found in newspapers and literature. The Eldridge list is probably the best known of this type.
2. Those made from the misspellings of school children in their daily work. These have usually been combined with lists selected by teachers in the various grades as words which children in these particular grades need to know. Perhaps the Boston and Kansas City public school lists are the best known in this line.
3. The words which children use in written composition. The one well-known study in this line is that of W. Franklin Jones out of which he secured his famous one hundred "Demons".
4. The frequency of usage in adult correspondence has been chosen as representing the real needs outside the school. The studies of Anderson and Horn are most complete in this field.
5. Perhaps a fifth might be added, that is, the list arranged by Dr. Starch of the University of Wisconsin which attempts to cover the whole range of non-technical English words.

Some argument may be presented in behalf of each of these types of word lists. From the standpoint of the public school which takes all the children of all the people, it is believed that the fourth form,

the standard of requirement in afterschool life, is the best. I do not mean to state that it is not desirable for children to learn to spell words which have but infrequent usage outside the schoolroom. I only maintain that the minimum list which every child should learn to spell to the point of automatic accuracy should be those words for which we have the best guarantee that the learner will have personal need during the years of life beyond the school age. For this reason, the major part of the discussion during this period will be based upon findings on words taken from correspondence.

Methods of Teaching and Learning. A considerable amount of experimentation has been conducted by psychologists and students in experimental education in endeavoring to find the most efficient method of teaching. Different methods of presentation, different time-intervals, and a multitude of devices have been originated and tried out with this purpose in mind.

Information has been secured from approximately one thousand teachers on the method used in teaching a spelling lesson. I have not found time to digest completely the material, but a somewhat cursory examination of the data indicates that there is truly a multitude of different ways in vogue.

The one thing that stands out clearest is that every child in the entire class is expected to learn each word in the assigned lesson by means of a particular teaching method or device. Experimentation has quite clearly demonstrated that a considerable proportion of the children in a class already know the spelling of from 10 to 75 per cent of the words in a given spelling lesson before the teaching process begins. Certainly we cannot defend the practice of requiring all the children to go thru the mechanics of learning that which they already know. If the lesson is assigned for silent learning without teaching, the children will certainly not devote the scheduled amount of time to the lesson before them. Why should they? If they are under a disciplinarian who insists that each child in the class shall be studying a particular book and lesson at a particular time, the deadening influence can be easily imagined.

If the lesson is taught by the teacher to the class, that is, if all the class goes thru the work in concert, which seems to be the prevailing custom, then the children are all participating, but some are participating in that which does not contain new, enlightening material and in which mental growth does not take place. Consequently, it seems that in spelling we should first test over the matter in hand and then have each child work upon the words which he himself missed. The possible detriment due to writing an incorrect form is more than overcome by the crystallization of motive in recognition of the specific task to be done and the specific goal to be attained. Experimental evidence thoroly supports this view.

I might say parenthetically that the same thing holds true in other subjects than spelling. One of the most enlightening procedures which I have known to be conducted in a public school system for a long time occurred in a school in which the supervising principal at the be-

gining of each six-weeks' period had the teachers give to the pupils an examination covering the points which they would be expected to know at the end of this period. The amount of material which was already known, and consequently did not need to be taught, was surprising. Much of it was known well, not only by a few bright ones in the class but by every member of the class.

On the basis of rather elaborate experimentation it seems that the important thing is not method of teaching but method of learning. What is needed is not a precise method of teaching lesson after lesson, but a clear, definite, usable method of learning on the part of the pupils. This method of attacking the learning of a word; of self checking, in order to know when it has been learned, is the most valuable thing which we can give our children. This has been quite thoroly worked out, at least a method of learning has been devised which gives very satisfactory results. The following will make clear the evidence that the method works:

RECORD OF TESTS

	A	B	C	D	1	2	3	4
condition					1	0	0	0
opinion		/			6	3	2	0
certain		/			1	2	1	0
piece					6	8	0	0
circular		\		+	9	5	2	2
organize					4	1	2	3
associate	—		—		10	6	2	0
business	+				11	6	2	1
<hr/>								
first	20	7	3	22	202			
second	11	1	1	4		123		
third	3	3	0	0			48	
fourth	1	2	0	0				24

In this chart A, B, C, and D represent different individuals in the class. 1, 2, 3, and 4 on the same line indicate the different tests over the material. But 8 words out of the 50 that were given the class to learn are shown on the chart. The figures under the numbers of the test opposite each word show the number of children in the class who misspelled the word on the test numbered at the top, while the numbers at the bottom of the chart after the words "first", "second", "third", and "fourth" respectively, show the number of words missed by the child indicated by the letter above. The totals given in the diagonal show the total number of words missed by all the children on each of the four tests.

The symbols opposite each word for the various individuals A, B, C, and D are interpreted as follows: A vertical line (|) means that the word was missed on the first test; a horizontal (—) that it was missed on the second test; oblique to the left (/) that it was missed on the third test; and oblique to the right (\) that it was missed on the fourth test. The uniting of any two of these symbols would indicate that this individual missed this particular word on the tests indicated by the symbols.

The chart is read thus: The word "condition" was missed on the first test by pupil "D", the vertical line showing that it was Test No. 1. Also the 1 in column for Test 1 shows that but one child missed this word on the first test and no child missed it on the succeeding tests. The word "opinion" was missed on the first test by child "A", as shown by the vertical line; it was missed on the third test by child "B", as is shown by the oblique to the left; and it was also missed on the first test by child "D", as is shown by the vertical symbol. The rest of the chart is interpreted in the same way. The totals show that 6 children missed it on the first test, 3 on the second, 2 on the third, and no one on the fourth.

It will be noted that the word "circular" was missed by child "D" on both the first and second tests as indicated by the vertical and horizontal symbols.

Dropping down to the bottom of the table you will note that child "A" missed 20 words on the first test, 11 on the second, 3 on the third, and 1 on the fourth. That is, he cleaned up 19 of the 20 words which he did not know.

Child "D" missed 22 on the first test, 4 on the second, and none on the third and fourth. The total sums show that 202 words out of the total possible number (50 words multiplied by 19, the number of pupils in the class), were missed on the first test; 123 were missed on the second; 48 on the third, and 24 on the fourth. That is, the class had cleaned up their spelling work from an average of almost 11 words missed per pupil to an average of 1.3 per pupil.

Measuring Spelling Results. The desire to be able to measure accurately the results of experimentation and the status and progress of the pupils in school has been responsible for the development of standardized tests and scales. Naturally, spelling was one of the early subjects in which this type of work was done. The Buckingham Scale which appeared in 1913, and the Ayres Scale in 1915, with the various tests which have been built, especially upon the latter, evidence the activity in this line.

The Ayres Scale has been of immense service to the schools of the country. It has focused attention upon a small list of very common words and has given, in a general way, an easy method of comparison of local results with general school results in the cities of the United States. Judged by this scale, progress from grade to grade is a very regular matter in the subject of spelling and the learning of words is a very uniform process. The naïve user of the scale, observing that words in column "J", for example, are listed as 66 per cent words for Grade II, 84 per cent words for Grade III, 94 per cent words for Grade IV, 98 per cent words for Grade V, and 100 per cent words for Grade VI, expects each word within that column to behave in just this manner. Experimentally it is discovered that they do not; that on the basis of averages for these words with a large number of children, these results are approximately correct, but individual words within this column may have varied somewhat widely. The following table shows this very clearly:

GRADE	II	III	IV	V	VI
south	36	65	96	98	98
game	56	75	91	99	98
age	68	86	95	99	100
band	62	77	91	91	97
glad	64	75	84	96	97
think	73	90	98	99	99
bill	62	73	95	98	100
gold	77	90	96	98	100
cannot	75	88	94	95	98
wind	74	82	97	95	97
girl	73	78	91	95	99
May	72	81	99	99	99
Average	66	80	94	97	98 ¹ / ₂
Ayres' Standard	66	84	94	98	100
Ayres' Variation	63-69	82-86	94-95	98	100
Variation above	36-77	65-90	84-98	91-99	97-100

This table shows the results of 200 children in each grade, the children being chosen at random from 25 to 40 Iowa cities. It is to be read in this manner: The word "south", children in Grade II spelled with 36 per cent accuracy; Grade III, 65 per cent; Grade IV, 96 per cent, Grade V, 98 per cent; and Grade VI, 98 per cent. The word "game", children in Grade II spelled with 56 per cent accuracy; Grade III, 75 per cent; Grade IV, 91 per cent, etc. The figures for the other words are read in the same manner.

These words were not chosen to present the most extreme variation. I am quite sure that I can find in my data words which appear in column "J" of the Ayres Scale which second-grade children spelled with a lower accuracy than 36, and other words which they spelled with a higher accuracy than 73. I selected as many as I thought could be well shown on a single chart showing some variation away from 66 per cent, but still averaging 66 per cent for Grade II children. The percentages for the other grades follow in order wholly without any thought concerning their selection.

You will note that the average in Grade II for these words as found in my data and on the Ayres Standard is identical. They were so chosen. In the other grades there is not very much variation. On the basis of the scale the third grade with 80 per cent is 2 per cent below the lower variation which Ayres included in his 84 per cent step, the ranges for these steps being shown as the "Ayres variation". By this is meant that in the making of his scale all accuracies of from 63 per cent to 69 per cent were included together as 66 per cent. Accuracies 82 per cent to 86 per cent inclusive were included as 84 per cent. In the last three steps of the scale but a single per cent is included, 98 per cent, 99 per cent, and 100 per cent each being a separate step. Thus the average for the fifth grade is one step low and for the sixth grade is 1½ steps. This is a situation which is found with a high degree of regularity, viz: that when the expectancy as per grade-steps falls in the extreme upper range, the tendency is for the actual results to fall short. Otis found

this same fact true, which he reported in *School and Society*, Vol. IV, in 1916.

More significant, however, than these slight variations in average are the variations in the individual words. While the average for even this small number of words brings the result very close to the results which Ayres places in his scale, the actual variation for individual words is much wider. This extreme is shown by the lowest figures on the table. "South" and "gold" both appear in the Ayres list, column "J", and the data which I have collected show that second-grade children spell "gold" more than twice as accurately as "south", and third-grade children more than 25 per cent better. This variation is shown in another way on the following charts. These charts show by steps the progress which one might expect from the position of the words in the Ayres scale and the actual progress which is shown by this study. (Charts not included.)

It will be noted that some of the words follow the expectancy quite closely while others differ considerably from it. This can be shown much more clearly on some other words than these under discussion. This fact presents a very interesting situation in the whole spelling problem. Some words seem to present great difficulty for children in two or more consecutive grades and then are suddenly learned to a rather high percentage of accuracy. Other words are learned to a rather high degree of perfection in the first two or three grades but continue to cause about so much trouble for the children in the various grades thereafter. The following illustrates this very clearly:

GRADE	II	III	IV	V	VI	VII	VIII
above	18	52	81	94	97	99	100
ache	8	13	37	65	83	85	88
aid	21	27	54	69	80	92	98
also	35	63	85	94	96	99	99
April	18	61	83	92	92	98	99
basket	28	57	88	92	97	97	98
bottle	16	49	74	89	96	98	100
branch	26	53	75	91	97	98	98
brought	28	54	85	90	94	96	98
butter	45	80	97	97	99	99	100
city	51	72	96	97	99	98	100
collect	6	8	29	64	84	93	96
colored	21	28	65	66	79	84	89
comply	15	37	60	79	85	88	88
exception	1	8	28	49	72	78	86
else	17	30	52	76	86	94	98
carrying	27	33	44	53	70	70	86
cross	37	60	80	94	98	99	100
rule	17	48	74	93	95	100	100
face	45	78	95	96	99	99	100
drop	48	69	88	92	97	99	99
lady	45	68	92	94	98	99	99
corner	44	61	70	80	90	92	95

The conclusion in this matter seems to me to be perfectly clear, namely, that when testing children with words from a column in a scale, it is necessary to recognize the following facts:

1. That the average for a large number of children, even upon a relatively small number of words, will approximate closely a satisfactory measure of the situation.

2. The larger the number of children that one is measuring, the smaller the number of words one may use and secure the same degree of accuracy.

3. In measuring as small a group as a class it is probably necessary to use as many as 50 words, and greater accuracy will be secured by using 100 or more.

4. Individual children on any group of words, and individual words with any size group of children, are likely to vary far from a stated norm.

Another development that has been receiving considerable attention in recent years is the question of trouble spots, places of special difficulty, in a word. Many people have assumed and frequently stated that every word has its particular place of difficulty. On this assumption it has been concluded that if only this trouble spot for each word were ascertained the whole difficulty of teaching spelling would be eliminated. Further, it has been assumed that these trouble spots will classify into a few types so that by having our attention called to a few types of prevalent error we may guard ourselves against misspelling of all words. Experimentation has not revealed such a situation. Experimentation has revealed that Jones' "100 Demons" are not the 100 demons for all children or for children in successive grades. Jones mixed together his findings in all the grades and thereby covered up some significant facts.

Furthermore, a checking of the spelling errors on a random selection of words has shown rather conclusively that many words do not exhibit clearly a specific type of error, while others do. Further, that the specific point of difficulty upon some of these words which tend to show such, changes from grade to grade. The following words will illustrate some of these findings, and should be read as follows:

Out of 200 attempts the fourth-grade children missed the word "extreme" 170 times, misspelling it in 59 different ways. The sixth-grade children missed the word 107 times in 29 different ways, and the eighth-grade children missed it 44 times in 10 different ways. The most frequent form of misspelling was "extream" which form occurred 43 times in fourth grade, 44 times in sixth grade, and 20 times in eighth. The next form "exstream" occurred 34 times in fourth, 10 times in sixth, and 1 time in eighth. "Extreem" occurred 7 times in fourth, 13 in sixth, and 11 in eighth. The other illustrative words which I shall give are all read in the same way. It is to be remembered that in each grade there were 200 children and that the numerator of the fraction gives the total number of times the word was misspelled and the denominator the total number of different ways of misspelling.

GRADE	IV	VI	VIII		IV	VI	VIII
extreme	170	107	44	extream	43	44	20
	59	29	10	exstream	34	10	1
				extreem	7	13	11

This word tends to show two types of errors in each grade, but while one type persists thru the grade, the misspelling which stands second in frequency in the fourth grade almost entirely disappears in the eighth.

The word "forenoon" is perhaps as good an illustration of a word with specific difficulty as I have been able to find.

GRADE	IV	VI	VIII		IV	VI	VIII
forenoon	118	68	39	fornoon	76	62	35
	27	6	5	furnoon	5	2	1

With 76 of the 118 errors in the fourth grade, 62 of the 68 in the sixth grade, and 35 of the 39 in the eighth grade, the omission of the "e" is certainly characteristic of the misspellings of the word.

On the other hand, the word "picture" presents a different situation:

GRADE	IV	VI	VIII		IV	VI	VIII
picture	55	14	3	pitcher	6	2	1
	34	11	3	pitchure	6	1	1
				pitcure	5	1	1
				piture	3	3	0

Out of 55 misspellings in the fourth grade, there were but six occurrences of a single type. Of 14 misspellings in the sixth grade, there were but 3 of a single type, and this was not the type which had 6 occurrences in the fourth grade. The word was misspelled but three times in the eighth grade but in three different ways.

The question may well be raised as to whether the type of error is of sufficient importance to be specifically pointed out if its frequency of occurrence among the total number of misspellings by 200 children represents less than 50 per cent or often less than 25 per cent of the total number.

"Publication" presents another good illustration of the diversified possibilities.

GRADE	IV	VI	VIII		IV	VI	VIII
publication	137	57	21	puplication	4	5	2
	126	48	19	publycation	4	1	0

Certainly there is no place in that word upon which one may put his finger and say "This is the trouble spot in learning to spell this word."

What shall we say about the word "recommend"?

GRADE	IV	VI	VIII		IV	VI	VIII
recommend	187	154	89	recomend	25	70	39
	109	44	11	recamend	12	9	3
				reccomend	5	16	30

Numerous other illustrations might be cited.

There are certainly some words which are not homonyms which yet sound so nearly like other words that they are frequently misspelled in spelling lessons.

The most prevalent way of misspelling "weigh" in the sixth and eighth grades was "weight". This occurred in more than 50 per cent of the total misspellings.

Many other illustrations of this type could be given. I think probably one would be doing a very distinct service in spelling, at least for those who teach by the usual method, if these words which children frequently confuse were to be carefully identified so that teachers would consciously make a special effort to pronounce these words distinctly.

Perhaps no particular type of word has received more attention from this standpoint than the *ei-ie* words. The most elaborate experiment conducted to ascertain the value of a rule in the learning of spelling was conducted by a University of Iowa student on this particular point. The data were gathered by means of a very carefully controlled teaching experiment on over three thousand children in three cities. The results showed that while the rule is more effective than drill, in improving the spelling of words not taught but governed by the rule, drill was more effective in correcting inversion errors; did not result in causing inversion errors, and was more effective in spelling words not governed by the rule and in securing permanent results.

Less elaborate experiments with other rules have tended to verify the same conclusion.

The recent developments in the field of spelling may be summarized, therefore, about as follows:

1. The consensus of opinion tends toward a list of words of greatest frequency in adult correspondence as the minimum list which all children in the elementary grades should learn to spell to a point of automatic accuracy. Other words of high social usage, or of great importance in school work, may be added as supplementary to this minimum requirement.

2. Since almost any list of words appropriate for a given grade contains words whose spelling is already known by a portion of the children in the grade, a preliminary testing should be given in order that time shall not be wasted in teaching children words which they already know.

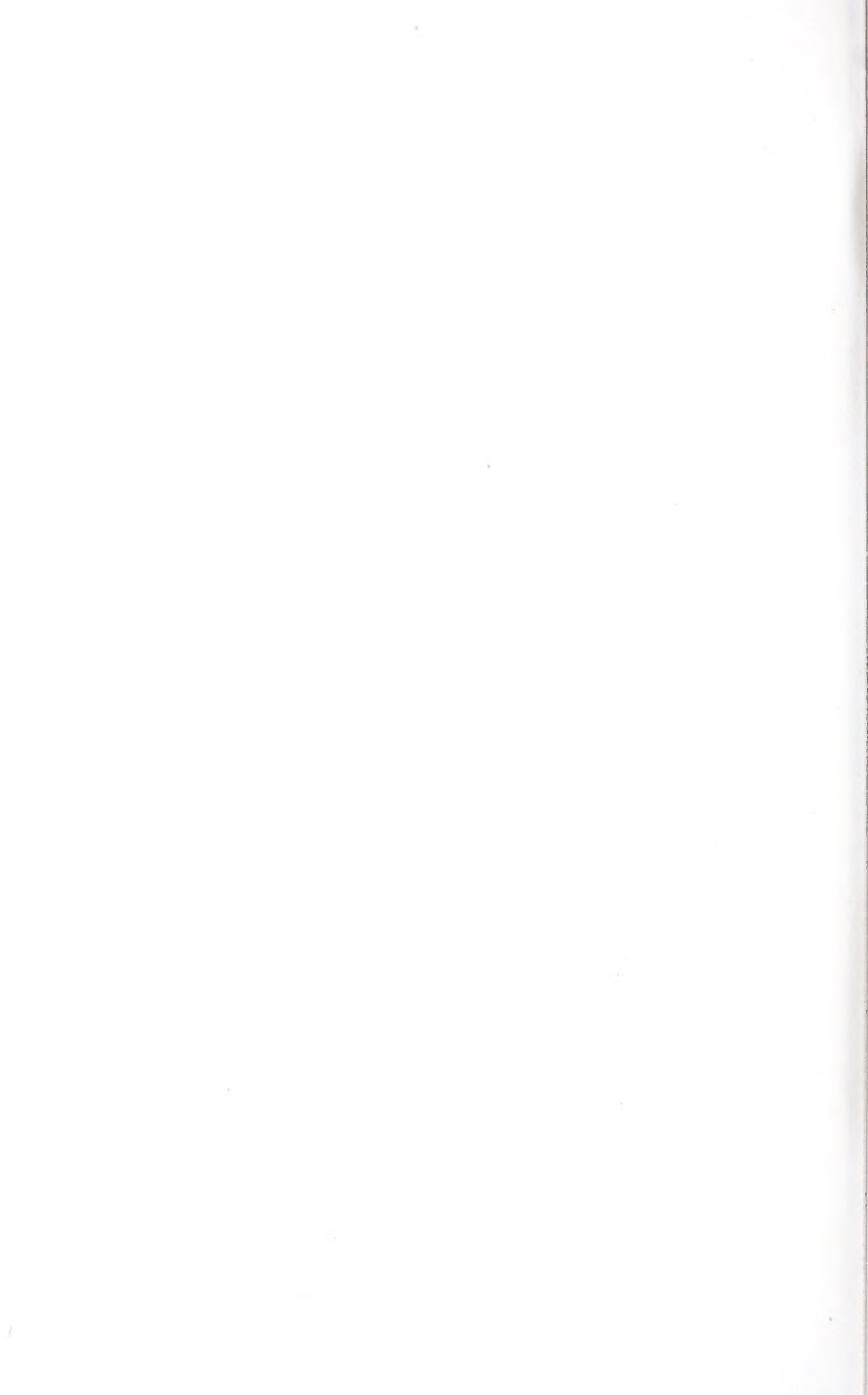
3. It is more important to give a child a definite usable method of learning, a tool which he may use in attacking his own individual problems, than in developing a high technique of teaching on the part of the teacher.

4. In measuring results of spelling, standardized tests and scales are of very great value. It must be remembered, however, that the standards on these tests and scales are mass results, and consequently, if a small list of words is to be used, a large number of children must be considered and vice versa, in order to have reliable results. Even then, it must be recognized that individual words and individual pupils may be expected to vary widely from a standard.

5. Trouble spots, the particular points of difficulty in words, occur in some words and do not occur in others. Further, the particular type of difficulty which seems prominent in one grade occasionally disappears and another form of misspelling becomes prominent. For this reason it is doubtful if it is worth while to spend much time in calling attention to trouble spots.

6. The advisability of teaching spelling rules has not been approved by the results of experimentation.

The points discussed in this paper, as well as many others, are discussed in Part 2 of the Eighteenth *Yearbook* of the National Society for the Study of Education. It was thought worth while, however, to present to you at this time experimental data on several of the points under discussion.



PUBLICATIONS OF THE BUREAU OF COÖPERATIVE RESEARCH

Requests for these publications should be addressed to the Bureau of Coöperative Research.

REPORTS OF CONFERENCES ON EDUCATIONAL MEASUREMENTS

- Proceedings of a Conference on Educational Measurements, April 17-18, 1914. Vol. XII, No. 10, *Indiana University Bulletin*. (Out of print.)
- Second Annual Conference on Educational Measurements, April 16-17, 1915. Vol. XIII, No. 11, *Indiana University Bulletin*. Price, 50 cents.
- Third Conference on Educational Measurements, April 14-15, 1916. Vol. II, No. 6, *Bulletin of the Extension Division*, Indiana University. Price, 50 cents.
- Fourth Conference on Educational Measurements, April 20-21, 1917. Vol. III, No. 8, *Bulletin of the Extension Division*, Indiana University. Price, 50 cents.
- Fifth Conference on Educational Measurements, April 19-20, 1918. Vol. IV, No. 4, *Bulletin of the Extension Division*, Indiana University. Price, 50 cents.
- Sixth Conference on Educational Measurements, April 18-19, 1919. Vol. V, No. 1, *Bulletin of the Extension Division*, Indiana University. Price, 50 cents.

STUDIES IN EDUCATIONAL RESEARCH

- Arithmetic: A Coöperative Study in Educational Measurements. By M. E. Haggerty. Vol. XII, No. 18, *Indiana University Bulletin*, *Indiana University Studies* No. 27. Price, 50 cents.
- Studies in Arithmetic. Edited by M. E. Haggerty. Vol. III, No. 32, *Indiana University Studies*. Price, 25 cents.
- The Ability to Read: Its Measurement and Some Factors Conditioning It. By M. E. Haggerty. Vol. IV, No. 34, *Indiana University Studies*. Price, 25 cents.
- A Coöperative Study of Reading in Sixteen Cities of Indiana. By Bureau of Coöperative Research. Vol. V, No. 37, *Indiana University Studies*. Price, 35 cents.
- Studies in Arithmetic, 1916-17. Edited by Walter S. Monroe. *Indiana University Studies* No. 38. Price, 35 cents.
- Progress and Promotion of Pupils in Certain Indiana City and Rural Schools. By Walter S. Monroe. *Indiana University Studies* No. 39. Price, 35 cents.